July 24/55

ELEMENTARY ECONOMICS

THE MACMILLAN COMPANY
NEW YORK · BOSTON · CHICAGO · DALLAS
ATLANTA · SAN FRANCISCO

MACMILLAN & CO., LIMITED
LONDON · BOMBAY · CALCUTTA
MELBOURNE

THE MACMILLAN CO. OF CANADA, LTD.
TORONTO

ELEMENTARY ECONOMICS

BY

FRED ROGERS FAIRCHILD

PROFESSOR OF POLITICAL ECONOMY

EDGAR STEVENSON FURNISS

PELATIAH PERIT PROFESSOR OF POLITICAL AND SOCIAL SCIENCE

NORMAN SYDNEY BUCK

ASSISTANT PROFESSOR OF POLITICAL ECONOMY

YALE UNIVERSITY

VOLUME II

New York

THE MACMILLAN COMPANY

1927

All rights reserved

PRINTED IN THE UNITED STATES OF AMERICA

COPYRIGHT, 1925 AND 1926,
BY FRED ROGERS FAIRCHILD.

COPYRIGHT, 1926,
BY THE MACMILLAN COMPANY.

Set up and electrotyped. Published May, 1926. Reprinted August, December, 1926; May, 1927.

Norwood Press
J. S. Cushing Co. — Berwick & Smith Co.
Norwood, Mass., U.S.A.

29-40

CONTENTS

VOLUME II

PART IV. PROBLEMS OF THE BUSINESS ORGANIZATION

PAGE

CHAPTER XXIX. THE ECONOMICS OF RAILROAD TRANSPORTATION . 3

The railroads and the public. Dependence of the railroads upon government. Public dependence on the railroads. Public utilities. The railway problem. Railway expenses. Fixed costs. Operating costs. Decreasing costs. Joint costs. Railway rates. Charging what the traffic will bear. Discrimination between goods. Railroad competition. Rate discrimination. Local discrimination. Examples. The reason for local discrimination. Another example. Conclusion. Personal discrimination. The breakdown of competition. Monopoly inevitable and economical. Regulated monopoly.

CHAPTER XXX. GOVERNMENT REGULATION OF RAILROADS IN THE UNITED STATES 26

State regulation of rates. The movement for strict regulation. The "Granger laws." The conflict of jurisdiction. Federal regulation: the Interstate Commerce Act of 1887. Defects of the Interstate Commerce Act. Development of regulation to the period of war control. Maintenance of competition. Limitation of rate increases. War time operation by the government: the railway situation. Government operation. The return to private operation. The Transportation Act of 1920. Provision for return of the roads. Financial provisions. Labor provisions. Railway consolidation. Railway securities. Railway service. Regulation on trial.

CHAPTER XXXI. INDUSTRIAL MONOPOLY AND ITS CONTROL . . 46

Definition and classification. Public legal monopolies. Private legal monopolies. Patents. The copyright. The exclusive franchise. Natural monopolies. Capitalistic monopoly. The foundation of capitalistic monopoly. The problem of capitalistic monopoly. The trust movement. Monopoly devices: the pool. The trust. The security holding company. Consolidation. Advantages and disadvantages of holding companies. Anti-trust legislation in the United States. The Sherman Anti-Trust Act of 1890. The operation of the Sherman Act. The anti-trust legislation of 1914. The Trade Commission Act. Investigation of corporations. Advisory functions.

PAGE

Unfair competition. The Clayton Act. Remedies provided by the Clayton Act. The Webb-Pomerene Act. Conclusion. Is monopoly inevitable? Forces favoring monopoly. Limits to monopoly.

CHAPTER XXXII. RISK, INSURANCE, AND SPECULATION . . . 71

Futurity, human knowledge, and risk. The human attitude toward risk. Risk reduced by knowledge. Preventive measures. Combining risks: the law of large numbers. Reserves. Passing risks on. Specialists in risk bearing: the entrepreneurs. Insurance. Insurance reduces risk. Insurance may reduce loss. Speculation. The technique of speculation. Speculation equalizes prices. Speculation promotes economical consumption. Usefulness of the speculative market. Hedging. Social service of speculation. Investment, speculation, and gambling. The corner. Judgment of speculation.

PART V. THE DISTRIBUTION OF WEALTH AND INCOME

CHAPTER XXXIII. ECONOMIC RENT IN AGRICULTURE . . . 97

The problem of distribution. The technical terms of distribution. Economic rent an income from land. Definition of economic rent. Economic rent equated to the productivity of land. The principle of diminishing returns. Diminishing returns in agriculture. What is implied in diminishing returns. Validity of the law of diminishing returns. Diminishing returns translated into increasing marginal cost. Marginal cost equated to market price. Producer's surplus drawn from the difference between average and marginal costs. Contractual economic rent equivalent to producer's surplus.

CHAPTER XXXIV. ECONOMIC RENT IN AGRICULTURE (*Continued*) . 114

Cultivation of lands of different productivity. The appearance of economic rent. Why production is pushed to the limit. Intensive and extensive margins of cultivation. Practical illustrations of the intensive margin. The extensive margin in practice. Economic rent equivalent to the contribution of land to the social income. Relation of economic rent to private property in land. Economic rent a regulator of agricultural production. Some social consequences of diminishing returns. Economic rent not a cost of production. Rent and cost of production in practice.

CHAPTER XXXV. RENT OF URBAN SITES AND MINES. PROBLEMS OF RENT 134

Producer's surplus or economic rent on urban lands. Diminishing returns from urban lands. Diminishing returns and the margins of utilization. Urban site rents in other uses. Comparative economic rent a governing force in the location of urban enterprises. Apparent exceptions. Economic rent not a cost-governing price. High rent

PAGE

and high prices. High rent and low prices. Rent of mines. Apparent differences between mines and other natural resources. Diminishing returns a principle of general applicability. Features which distinguish rent from other income. Attacks on private receipt of rent. The problem raised most clearly by urban site rents. Large visible rents. Absentee ownership. Influence of situation obvious. The concept of unearned increment. The Single Tax movement. The case for the Single Tax. The case against the Single Tax. Appropriating future rent increments.

CHAPTER XXXVI. THE NATURE OF INTEREST. THE SUPPLY OF LOANS 163

Land and capital: landowners and capitalists. Production of capital requires time. Time preference. Interest the reward for waiting. The universality of interest. The form and measurement of interest. Essence of the loan contract. Interest not caused by the use of money. The loan market and the rate of interest. The supply of loans. Marginal utility in present and future. Time preference again. The supply curve. Character of individual saving. Involuntary saving. Corporate saving. Saving by the insurance companies.

CHAPTER XXXVII. THE DEMAND FOR LOANS. THE RATE OF INTEREST. THE SOCIAL PROBLEM OF INTEREST . . . 177

Borrowing for production. The productivity of capital. Diminishing productivity from capital. Borrowing for consumption. Individual time-preference. Marginal time-preference. Borrowing of governments. Total demand for loans. Determination of the market rate of interest. Relations of the individual to the loan market. The rate of interest and the quantity of money. Transition effects. Interest rate a factor in all valuations of income. Lack of uniformity in interest rates. True and nominal interest rates. Changes in the rate of interest. Future of the interest rate. The capitalist class. Opposition to interest as a private income. Early regulation of interest. The modern field for regulation. Futile and injurious interference with the interest rate.

CHAPTER XXXVIII. THE GENERAL LAW OF WAGES . . . 202

The rewards of human exertion. The problem of the rate of wages. The demand for labor. Diminishing productivity as applied to labor. Law governing an employer's demand for labor. Wages are advance payments. The general demand for labor. The supply of labor. Demand and supply; the market rate of wages. Influence of the supply of labor. Relation of the individual employer to the labor market. Practical qualifications. Dependence of different types of labor. The sphere of bargaining. Wages above marginal product. Wages below marginal product. Lack of uniformity in wage rates. Forces tending toward uniformity: machinery. Trade unionism.

PAGE

Systems of wage payment. Piece rate wages. Piece rates and the economic law of wages. Variations in local wage rates. Marginal productivity of fundamental importance. Other types of labor income.

CHAPTER XXXIX. DIFFERENCES IN WAGES. WAGE PROBLEMS . 231

The principal labor groups. Wages of each group tend toward uniformity. Least uniformity in the highest group. Wages and the contrast of social classes. The cause of discrepancies between wages of the different classes. Permanence of the cleavage between social classes. Nature and nurture theories. Social classes chiefly self-perpetuating. Wage problems. Real wages. Long run changes in the prosperity of labor. The law of proportionality. Historical changes in the rate of wages. Law of proportionality applied to specific labor groups. The standard of living theory of wages. The true effect of the standard of living. Elevation of the wage rate through action of the employer. Minimum wage laws.

CHAPTER XL. PROFITS 256

An example of profits. Rent and interest are costs. Wages of management a cost. What governs the rate of wages of management? The definition of profits. The existence of profits. Losses. The balance of profit and loss. Profits a reward for risk-bearing. Risk, insurance, and profits. The promoter and his profits. The small scale risk bearers. Not all risks inevitable. Profits tend to be short-lived. Profits and monopoly. Profits not a price determinant. Receivers of profits control modern industry.

CHAPTER XLI. PERSONAL DISTRIBUTION 273

Functional shares in distribution. The total money income of the people of the United States. Total income corrected for price changes. Comparative national incomes. Inequality of personal incomes in the United States. Inequality universal. Inequality graphically presented. Personal distribution of wealth. Evidence from American probate courts. Personal distribution of wealth in other countries. Labor's share of the national income. The problem of poverty. Family incomes and living costs. Minimum standard of living in America. American family incomes. Qualifications and conclusions. Problems raised by inequality. Is inequality increasing? Growing inequality in America. Evidence from other countries. Causes of inequality. Personal ability and inequality. Inheritance and inequality. Reform of inequality. Charitable relief. Relief from the hazards of life. Public education. Limitations upon individual freedom. Restriction of property rights.

CHAPTER XLII. POPULATION 299

The problem of population. Malthusian theory of population. Criticism of Malthus' theory. Modified statement of the Malthusian

PAGE

theory. Nineteenth century expansion. Over-population still a menace. How fast can population grow? Checks to population increase. Positive checks. Preventive checks. The death rate. The birth rate. Causes of the declining birth rate. The differential birth rate. The question of quality of population. Eugenics.

CHAPTER XLIII. IMMIGRATION 319

Spread of population. Immigration to the United States. Causes of the changed character of immigration. Restriction. Exclusion. Selection. Restriction. Restriction acts. Consideration in favor of the act of 1924. Arguments in opposition to the act of 1924. Social consequences of the restrictive policy.

PART VI. THE ECONOMICS OF GOVERNMENT

CHAPTER XLIV. PUBLIC FINANCE 341

Government expenditures. Analysis of public expenditures. The cost of war. Other national expenditures. State and local expenditures. The growth of public expenditures. Total cost of government in the United States. The cost of government services. Government revenues. Industrial earnings. The finances of government industry. Financing the primary functions of government. Taxes. Fees. Special assessment. Taxes, fees, and industrial earnings compared. Names may be misleading. Relation between expenditure and income. Borrowing to meet a current deficit. Emergency borrowing: war. Borrowing for investment or public works. The nature of government debts. The security back of public debts. The high standing of government obligations. The national debt of the United States. The World War debts. Government surplus. The budget.

CHAPTER XLV. TAXATION: GENERAL PRINCIPLES 365

How is the tax burden distributed? Equal distribution. Benefit received. Cost of service. Ability to pay. How is ability to pay measured? Proportional taxation. Progressive taxation. Regressive taxation. The tax-free minimum. Other measures of tax obligation: consumption taxes. Business taxes and poll taxes. All taxation cannot be apportioned according to ability. Escape from taxation. Tax evasion. The shifting of taxes. Direct and indirect taxes. The purpose of taxation. Some technical terms.

CHAPTER XLVI. THE GENERAL PROPERTY TAX 381

Evolution of the general property tax. In America. Typical organization and operation. Assessment. Appeal, revision, and equalization. Apportionment; the tax rate. Collection. Modifications. Position of the general property tax. Critical analysis of operation.

PAGE

Definitions. Assessment: discovery of taxable property. Valuation. Theoretical questions. The assessor's problem. Expert assessment the exception. Information from the owner. Where assessment is easiest. The attitude of the taxpayer. Undervaluation. How the assessor meets the situation. The results of assessment. Results of the breakdown of assessment. Inequality and the legal remedy. Theoretical weaknesses of the general property tax. Modifications in detail. Reform of the general property tax.

CHAPTER XLVII. THE UNITED STATES REVENUE SYSTEM. THE INCOME TAX. SHIFTING AND INCIDENCE OF INCOME AND PROPERTY TAXES 405

General property tax and income tax. Early prejudice against taxation in America. Constitutional restriction of direct taxation. Indirect taxation: customs and excise. The Civil War and the national tax system. Tariff and excise since the Civil War. Beginnings of the income tax. The income tax and the Constitution. The Sixteenth Amendment and the triumph of the income tax. The present revenue system. The income tax of 1926. Individual incomes. The tax rates. Earned incomes. Income of corporations. An example. The exempt minimum. What is taxable net income? Tax-free securities. The volume of tax-free securities. Escape of the wealthy taxpayers. What the government gains from tax-exemption. The effect of progression. State and municipal bonds and the Constitution. Salaries of public employees. Dividends of domestic corporations. Income tax rates. High rates due to war. Stoppage at source. Personal declaration. American practice. State income taxes. Shifting and incidence of taxes on property and income. The tax on economic rent. The tax on land values. The practical result. Capitalization of land taxes. Taxes on dwelling houses or on their rent. The tax on business property. The property tax on consumers' goods. The income tax.

CHAPTER XLVIII. THE CUSTOMS TARIFF. PROTECTION . . . 433

The customs tariff. Technical and administrative features. Shifting and incidence. Economic effects of the tariff. A numerical illustration. Conclusions as to incidence and economic effects. Two purposes of the tariff. The revenue tariff. The protective tariff. Fiscal weakness of the protective tariff. The controversy over protection. The case for protection. (1) The favorable balance of trade argument. (2) The gold argument. Gold and the national wealth. (3) The added industry argument. (4) The wages argument. Wages and labor cost. Effect of withdrawing protection. (5) The infant industry argument. What is the territorial unit for protection? Conclusion. Military necessity. Protection by bounties. Ship subsidies.

PAGE

CHAPTER XLIX. OTHER FORMS OF TAXATION 453

Excise taxes in the United States. Shifting and incidence. Effect on production. Elasticity of purchase. Technical and administrative features. Consumption taxes generally regressive. State excise taxes. Taxation of corporations. Corporations and the income tax. Difficulties with the property tax. Public utility corporations. Taxation of financial corporations. Exemption of securities. Special excise taxes on corporations. Business taxes. The inheritance tax. Typical technical features. Inheritance taxation by the American states. The federal estate tax. State and inter-state complications. Some fundamental questions. The poll tax.

CHAPTER L. GOVERNMENT INDUSTRY 469

The reasons for government industry. Government ownership in the United States. The United States Post Office. Postal rates. The flat rate. Rate discrimination. The postal deficit. The case for government ownership. Considerations against government ownership: interest and profits. Discrimination. Politics. The motive in business. General conclusion. The case of monopoly. Conclusion as to public utilities.

PART VII. PROBLEMS OF LABOR

CHAPTER LI. UNEMPLOYMENT 485

Origin of the labor problem. Extent of unemployment. The unemployed. The causes of involuntary unemployment. Invention. Seasonal fluctuations. The labor reserve. The social injury. Labor's loss through seasonal fluctuations. Effects of the business cycle. Defensive policies of labor: the full crew. Labor's proprietary interest in the job. The "lump of labor" theory. Remedies for unemployment. Adjustment of the labor supply. Peculiarities of labor supply: impossibility of storage. Rigidity of the wage rate. Immobility of labor. Labor exchanges. Private labor exchanges. Public labor exchanges. Stabilizing the demand for labor: public works. Stabilization by private enterprise. Unemployment insurance. The case for insurance. The Ghent system. Unemployment insurance in Great Britain. The weakness of public unemployment insurance. Unemployment insurance through private action.

CHAPTER LII. PROBLEMS OF THE WORKING DAY, CHILD AND WOMAN LABOR, AND INDUSTRIAL ACCIDENT 511

Modern aspects of the problem of the working day. Background of the movement toward shorter hours. Relation of hours to wages. Reduction of hours and increased productivity. Limits to reduction.

PAGE

Effects of invention. Modern aspects of child and woman labor. Extent of child labor. Effects of child labor. Legal regulation of child labor. In America. Extent and character of industrial labor of women in the United States. Causes of the increase in woman labor. Results of women's increased part in industry. The problem of equal pay. Underpayment of women. Legal regulation of woman labor. The waste of industrial accident. Attempts to reduce the waste of industry. The unavoidable loss of accident. Distribution of the loss. Damages under the common law. Compensation laws. History of compensation laws and insurance.

CHAPTER LIII. THE ORGANIZED LABOR MOVEMENT 535

Foundation of the labor movement. Origin of the labor movement in Great Britain. Conflict with the law. Recent developments in Great Britain. Early history in the United States. Adaptation of union structure to early economic conditions. Nation-wide organizations. Comprehensive organization of the wage class. Types of union structure: the trade union. The industrial union. Political activity of American unions. Obstacles to political action. Possibility of future political activity. Extent of the American labor movement. Union membership and the business cycle. Relative importance of organized labor. Where unionism does not flourish: among women workers. Among the unskilled workers. Non-economic activities of the unions. Benefit systems of the unions. Collective bargaining. Gains through collective bargaining. Conflict over the method of collective bargaining. The closed shop. Restrictive tendencies of union policy. The strike. Statistics of strikes. Violence in strikes.

CHAPTER LIV. SOLUTIONS OF THE INDUSTRIAL CONFLICT . . . 569

Protecting the public against strikes. Failure of legal prohibition of strikes in Great Britain. Legal status of the strike in America. Failure of the injunction. Conciliation. Conciliation promoted by government. Employee representation. American experience with employee representation. Mediation. Official mediation. Voluntary arbitration. Arbitration agreements and procedure. Official arbitration. Results of voluntary arbitration. Introduction of compulsion. Compulsory investigation. Appraisal of compulsory investigation. The case for compulsory arbitration. Experience in New Zealand and other states of Australasia. Compulsory arbitration in Kansas. Judgment of the Australasian experience. Inherent defect of compulsory arbitration. The opening wedge of socialism. The future of the industrial conflict.

PART VIII. REFORM OF THE ECONOMIC SYSTEM

PAGE

CHAPTER LV. PROFIT SHARING, CO-PARTNERSHIP, AND COÖPERATION 595

Profit sharing. Prerequisites of true profit sharing. Typical plan of profit sharing. Profit sharing may lead to producers' coöperation. Examples of profit sharing. Criticism of profit sharing. Limited usefulness in promoting efficiency. Limitations as an agency of industrial peace. Co-partnership. Examples of co-partnership. Progress of co-partnership. Weaknesses of co-partnership. Judgment of profit sharing and co-partnership. Coöperation. Producers' coöperation. Form of organization. Experience. Cause of failure. Reversion to the wage system. Credit coöperation. Schulze-Delitsch and Raiffeisen banks. Spread of the system. Consumers' coöperation. The Rochdale system. Principles of organization and management. Extent of the movement. Spread into allied fields. Conservatism of consumers' coöperation. Avoiding the wastes of competition. Broader social aspects. Social program in England. In other European countries. Obstacles to consumers' coöperation in the United States.

CHAPTER LVI. SOCIALISM 623

Breadth of the subject. Economic analysis of socialism. Definition. Socialist criticism of capitalism. The doctrines of Karl Marx. Materialistic interpretation of history. The class struggle. Marxian economics. Laws of capitalist development. Tendency toward concentration. Increasing number and misery of the laboring class. Socialist programs. Collectivism or state socialism. Syndicalism. Guild socialism. Communism. Failure of Russian communism. The soviet organization. The method of communism. Extent of the socialist movement. In Germany. In Central Europe. In Western Europe. In Great Britain. In the United States. Superficial criticism of socialism. Judgment of socialism.

INDEX 651

PART IV

PROBLEMS OF THE BUSINESS ORGANIZATION

ELEMENTARY ECONOMICS

CHAPTER XXIX

THE ECONOMICS OF RAILROAD TRANSPORTATION

The railroads and the public. Our previous study has shown how completely the modern system of division of labor is dependent upon adequate transportation facilities and how the railroads have gradually attained to the position of predominance in the transportation system. The railroads are the heart of the transportation problem.

In some countries railroad transportation is a government function. In the United States and many other countries however the railroads are owned and operated by corporations, whose stocks and bonds are the property of thousands of individual and corporate investors; they are private wealth, owned and operated for private gain.

Dependence of the railroads upon government. The fact of private ownership and operation however does not prevent the railroads standing in a class apart from the ordinary types of business enterprise, with public relations of the utmost importance. On the one side the railroads stand in a peculiar position of dependence upon the government. So great is the public interest in the development of transportation, that the government has in many ways come to the aid of the railways in their initial stages. We have seen in an earlier chapter (Chapter IX) that generous land grants were made by the federal government to several of the railroads and that in many cases state and local aid was forthcoming to assist in their building. What is of far greater importance, the government has given the railroads the advantage of

its right of *eminent domain*, by virtue of which private property can be condemned for public purposes at a price determined by the courts or other authority. The railroads have to acquire their "right of way" by the purchase of land from private owners, and despite the fact that in the pioneer days of railroad building the advantages which might accrue from proximity to a railroad line were so highly esteemed that in many regions the right of way was donated and in some cases even the grading of the right of way was done gratuitously by the people of the community, the power to secure the condemnation of land was in many instances of inestimable advantage to the railroads in enabling them to secure the most desirable routes at reasonable prices. The public has, in fact, entered into a quasi-partnership with the railroads in these respects.

Public dependence on the railroads. The dependence is not all on one side. The reason back of government aid is the supreme dependence of the people upon the railroads. A railroad is not a mere private business. The local department store can conduct its business about as it sees fit. It can sell what it chooses, refuse to sell certain things, or close up shop entirely; and it can charge whatever prices it can get. If the people do not like its goods or its prices, they are free to trade elsewhere. With the services of the railroads the situation is entirely different. The people are absolutely dependent upon the railroads, in most cases upon one railroad. This dependence in the United States has reached the point where a stoppage of service may assume the proportions of a national calamity and where inadequate service or extortionate rates may neutralize the advantages of large scale production or at any rate impose a great handicap upon producers and consumers.

Public utilities. In return for its coöperation the government therefore imposes upon the railroads the obligation to give adequate service to the public. The right of way and equipment must be kept up according to certain standards, a certain minimum number of trains, passenger and freight, must be run, and schedules must be maintained at all times, in winter and summer, in good weather and bad. Safety devices must be installed, and finally the rates charged for railroad services must not be unreasonable. The public

has become a partner in the railroad business, it has a vital stake in railroad services, and in return for the privileges granted it demands of the railroads adequate service and reasonable charges. This double relation of dependence and obligation between the railroads and the public is what puts them, along with certain other public service enterprises, such as the street railways, the telephone, the telegraph, water supply, etc., in a special class apart from private business in general, to which has been given the appropriate title of *public utilities* or *public service corporations*.

The railway problem. The foregoing paragraphs have served to indicate the public utility character of the railroads and what we may call the ideal relation of mutual dependence and obligation between the railroads and the public. Experience however has shown that this ideal is not realized automatically through the operation of economic forces, such for example as determine the quantity, the quality, and the prices of the products of competitive industry in general. Competition has failed to bring corresponding results in the field of railway transportation. When competition has prevailed, it has generally been of the cutthroat variety and has brought in its wake disturbing discriminations in rates, disorganization of service, and other abuses. On the other hand competition has frequently broken down altogether, introducing serious problems of monopoly. The investment in the railroads is so great that their methods of financing have attained public importance. Differences between the railroads and their employees are matters of vital public interest, since railway strikes or lockouts may easily play havoc with the nation's commerce. There is thus a "railroad problem," whose public importance can scarcely be exaggerated. Though, as we have seen, it has many sides, the heart of the problem is in the determination of the rates charged by the railroads; this chapter will therefore be devoted primarily to the investigation of railway rates. We shall find it convenient to posit the first steps of our inquiry on the assumption of free competition and the entire absence of governmental regulation, although the reader will doubtless recognize that this is not the situation existing today in the United States, whose railroad

conditions are the result of some forty years of regulation both by state and federal authorities.

Railway expenses. Railroading furnishes the most striking example of an industry subject to the law of decreasing costs. The manner in which this condition emerges in an industry where there is economy in large capital investment and in which the plant of maximum efficiency is a huge organization has been made sufficiently clear in an earlier chapter (Chapter XVII). The railroad business represents an extreme example of this.

Fixed costs. In the first place the railroads require an enormous investment of capital. Great expense must be incurred to bring a railroad into being. The right of way must be secured and improved, tracks must be laid, stations for freight and passengers must be erected, expensive equipment in the way of rolling stock, signalling apparatus, etc., must be purchased or leased. Funds for these purposes are secured through the sale of railway bonds or other securities; and the interest on bonds and notes, which must be paid at definite rates, as well as the funds which the railroad may have agreed to set aside each year to provide for the eventual redemption of such securities, taxes and the rentals of leased lines, form the greater part of the railroad's fixed charges. In harmony with the nature of fixed charges,[1] these expenses bear practically no relation to the volume of traffic. Given the facilities which are prerequisites to the satisfactory handling of any traffic, an increasing amount of goods and passengers may be hauled over the road without increasing the fixed charges, up to that point where it is necessary to enlarge the existing facilities — double-track the system, provide four tracks in place of two, build extra switching connections, enlarge the terminals, etc. — in order to take care of additional traffic. When this point is reached additions to the fixed charges may be expected, but again, once they have been incurred, they will not vary until it is again necessary to enlarge the plant.

It should be noted that the investment in a railroad is relatively

[1] See Chapter XVII, p. 321, for an analysis of fixed costs, general operating costs or overhead, and particular operating costs.

permanent, in the sense that there is almost no alternative use to which the equipment can be put if the railroad proves to have been an unprofitable venture. A factory building can sometimes be transformed into a storage warehouse, or diverted to the making of another article, but when once rails have been laid they can be used only to provide transportation, with the sole alternative of sale as scrap iron, and there is almost no way of recovering the very great expense of building the right of way, with its extensive cuts, grades, bridges, and tunnels. As always, the fixed charges are recurring costs, which represent the hire of capital funds which have been invested in the past, and which continue to accrue whether the road be operated or not. If the road should become bankrupt and be declared unable to meet these obligations, they are simply shifted once and for all from the railroad to the bondholders.

Operating costs. The operating costs, which at any time could be avoided by the decision to go out of business, may be subdivided into the expenses for the maintenance of way and structures, the maintenance of equipment, the cost of conducting transportation, and general expenses which cannot readily be allocated to any of the other groups. These expenses fall into the two groups of general operating costs — overhead and particular operating costs.

In the first group come most of the general expenses, of which the chief item is the cost of maintaining general offices for the railroad system, which cost bears but little relation to the size of the operations. The salaries of the officials go on regardless of the fluctuations in the volume of traffic, and the officials can handle a large volume of business with little extra expense, save for the clerical force. The cost of additional clerks, or the saving through the dismissal of clerks, is a comparatively small item in the total expense.

The maintenance of the right of way includes rail and tie replacement and supervision. While the rail maintenance is closely connected with the speed, weight, and number of trains, there is little relationship between the cost of renewing ties and the amount of the traffic. Tie replacement is necessary because of exposure

to the weather and would have to go on so long as any traffic is to be carried. Furthermore the expense of tie replacement forms a much greater fraction of the total expense of maintaining the right of way than does rail replacement. The cost of superintendence bears some relation to the volume of traffic, but it does not increase in proportion to increases of traffic.

The maintenance of equipment varies according to mileage, but not in proportion to the volume of goods carried. Full utilization of the carrying space of freight and passenger trains and of the tractive power of locomotives does not cause a proportional increase in the charges for repairs as compared with incomplete utilization. It is estimated that about one half of these expenses may be considered to be independent in large measure of the volume of traffic.

The cost of conducting transportation consists chiefly of the wages of the trainmen, fuel for the locomotives, and terminal charges. The wages and fuel charges vary in proportion to the train miles, but they do not vary directly, for all trains are not of the same length, nor are they equally heavily loaded. The wage cost for a full train is not much greater than for an empty train, and while the fuel cost increases with the load to be hauled, it does not vary proportionately. The terminal expenses vary with the volume of business, but bear no relationship to the train miles. Here again a large percentage of the expenses, perhaps one half, may be considered as independent of the volume of traffic.

Decreasing costs. We see therefore that not only are the fixed costs of the railroad business especially heavy as compared with the operating costs, but that of the latter group a relatively large part is assignable to overhead, leaving a comparatively small part of total railroad costs in the class of particular operating costs, which vary in accordance with the amount of business done. It has been estimated that about one half of the operating expenses and two thirds of the total costs can be considered as constant.[1]

The presence of this large proportion of constant expense indicates that a railway operates under conditions of decreasing costs,

[1] Eliot Jones, *Principles of Railway Transportation*, p. 78.

or increasing returns. Up to the point of maximum utilization, an increase in the number of passengers to be carried or in the amount of freight to be hauled does not cause a proportional increase in the total expenses. Hence the unit expense will decrease. If we assume that the rates remain constant, while the traffic increases, profits must show an increase which is proportionally greater than the increase in the volume of business, and the reverse would be true if the rates remained constant while the volume of traffic declined. This explains the fact, which has frequently been noticed, that a moderate growth in the volume of traffic, with no increase in the rates, is often enough to raise a railroad from financial embarrassment to the position of a highly profitable business enterprise.

This statement that the railroad is an industry of increasing returns may be given concreteness by an illustration. The figures in this illustration are chosen somewhat arbitrarily but they will serve to illuminate the point. "Assume that a railroad has a capital stock of $100,000,000, and an income of $100,000,000; and that its fixed charges amount to $28,000,000, and its operation expenses to $66,000,000.

Gross revenue		$100,000,000
Expenses		
Fixed charges	$28,000,000	
Operating expenses	66,000,000	
Total expenses		94,000,000
Profits (available for dividends)		$6,000,000

The railroad after meeting all its expenses has left the sum of $6,000,000 which equals 6 per cent on its capital stock.

"Assume now a 10 per cent increase in traffic. The revenue, on the assumption that the rate is the same on the additional traffic as on the old, will be $110,000,000. The fixed charges will be the same as before, it being assumed that the existing plant is adequate to handle the additional business. The operating expenses are not constant, as are the fixed charges; yet neither do they increase in the same proportion as the traffic. Instead of

increasing by 10 per cent, therefore, they increase by only 50 per cent of 10 per cent. The account therefore stands as follows:

Gross revenue		$110,000,000
Expenses		
Fixed charges	$28,000,000	
Operating expenses		
$66,000,000 + (50% of 10% of $66,000,000)	69,300,000	
Total expenses		97,300,000
Profits available for dividends		$12,700,000

The railroad, after meeting all its expenses, has left the sum of $12,700,000, which is equal to 12.7 per cent on its capital stock. An increase of 10 per cent in the traffic has thus increased the profits by over 100 per cent." [1]

Joint costs. Suppose it is desired to know the exact cost of shipping a ton of coal from Scranton to New York. We will assume that two freight trains, one consisting of twenty cars, the other of thirty cars of coal, are sent from Scranton to New York. It is quite possible to determine with a fair degree of accuracy the particular operating expenses of each train. There will be charges for fuel, wages for the trainmen, some wear and tear on the tracks and the rolling stock, and a few other expenses which might have been avoided entirely had the particular train not been run. These expenses will be greater for the second train than for the first, but they will not be fifty per cent greater, and if we were to determine the proportion of the particular operating expenses which should be borne by each ton of coal we would find the amount somewhat greater in the first case than in the second.

A third train of thirty cars is sent from Scranton to New York, but in this case it carries mixed freight — coal, cement, hardware, textiles, and groceries — and furthermore some of the cars are incompletely loaded. Again we can determine the particular running expenses for the train as a whole, but to distribute them among

[1] Eliot Jones, *Principles of Railroad Transportation*, pp. 78–79. Minor changes have been made in the illustration.

the different goods carried according to the expense caused by each individual shipment is another matter. The easy solution would be to calculate a flat rate according to weight or bulk, but it is quite clear that this does not accord with our desire to find out the expense of shipping each individual package. It costs more, for example, to handle a ton of small packages than a ton of coal. A car fully loaded with textiles costs little more to haul than one only half loaded, and a meticulous effort to allocate the particular operating expenses according to the cost of rendering the service might lead to the surprising result that it cost more to carry some of the textiles than others on the same train.

An accurate apportionment of the particular operating expenses presents enormous difficulties, but when we take the other expenses into consideration the case becomes hopeless. The fixed charges and the general operating expenses are incurred to provide equipment and personal services which are available not only for the passenger and freight services, but for specialized services within these two great groups, and which cannot be apportioned on any scientific basis. Such costs we call *joint costs*, meaning *costs which are incurred in the simultaneous rendering of two or more separate services and which can be apportioned only arbitrarily between the separate services or products*. With regard to the greater part of a railroad's expenses it is impossible to tell what portion should be assigned to passenger service and what portion to freight service, to say nothing of further assignment of costs within each of these services. It is just as impossible as to say, for example, what part of the cost of raising cotton should be attributed to the cotton fibre and what part to the cotton seed.

Railway rates. In competitive business generally, cost of production has an important if not controlling voice in the determination of price, and even where monopoly prevails cost generally at least sets a minimum below which price will not fall. In railroading, owing to the presence of joint costs, it is impossible to tell absolutely what any particular railroad service costs, although the additional out-of-pocket expense for affording an extra increment of service can be closely estimated. Specific cost can therefore

evidently not furnish the principle for determining rates. Any theory that rates "ought" to be adjusted to the cost of rendering specific services is therefore untenable. In practice the railroad managers know what their total costs are, and they seek to meet these total costs and earn what profit they may from the traffic as a whole.

Charging what the traffic will bear. In the absence of any scientific method of determining railway rates, the officials charged with the duty of rate making have recourse to a practical method, which has been given the somewhat repellent name of "charging what the traffic will bear." Some goods can stand a higher rate per ton than others. A ton of pig iron may be worth from $25 to $30, while a ton of silk cloth would be worth thousands of dollars. A freight rate per ton or per pound which would be prohibitively high if applied to pig iron might have little effect on the selling price of silks and consequently on the volume of traffic in silk. A high rate therefore is charged for carrying goods which combine high value and small bulk, while rates on cheap bulky goods are relatively low.

There are of course limits to the rates determined on the principle of what the traffic will bear. The upper limit is that rate at which shippers would not send the freight or passengers would not travel. The long run aspect of this should be emphasized, for a rate which would at the moment give the maximum return might be disastrous in the long run. It is not to the interest of the railroads either to cause the local industries to wither up and die or to force their relocation in places where more reasonable rates are charged. It is this limit which varies so widely according to the value and the bulk of the goods. The lower limit is the rate at which the traffic would yield a revenue to the railroad no greater than the additional cost which it would occasion. At this minimum rate, the traffic may not pay its share of the fixed or general operating costs. However, so long as it pays enough to cover its own particular operating costs (estimated as closely as the traffic officials are able) and ever so little above, it is good business to take it. So the cheap bulky goods, such as coal, trap rock, etc., are carried for

what is actually less than the average ton mile cost, while other parts of the service are counted upon to make up the fixed and general operating costs for the whole. In this way it is planned that the total revenue shall cover the total expenses and yield as large a net revenue as possible.

Discrimination between goods. Discrimination between goods is thus frankly an essential part of the prevailing method of determining railway rates. The word "discrimination" has an ugly sound, and a connotation of injustice is apt to attach to it. But to one who understands the nature of the railway business and of railway rates, it will be clear that such discrimination is not only inevitable but is in the interest of the people as a whole.

It will be granted that no railroad would be built unless the promoters of the venture could, with some degree of confidence, count on securing a gross revenue sufficiently large to cover all costs and yield a profit. If they are forbidden to charge different rates for transporting different classes of goods, it may not be worth their while to build the road, for a rate high enough to secure the maximum contribution toward the constant expenses from the high-grade freight would preclude the possibility of shipping bulky freight, while a rate low enough to attract the low-grade freight would materially reduce the profitableness of carrying the other. In other words, discrimination between goods may be the only condition under which a railroad can be built, and however much aggrieved a shipper of textiles may feel because he pays a higher rate per hundred pounds than a shipper of wheat, he is still better off than if there were no railroad, as his patronage of the road demonstrates.

Suppose now that a railroad, already built and in operation, sees an opportunity to add a new class of traffic, but only at a rate considerably below the scale paid on its present business. The new business will add little or nothing to the fixed and general operating expenses of the road, while the revenue obtained will contribute something, though perhaps not much, over and above the particular cost added by the new traffic. This will be profitable to the railroad, and it does no injury to the former patrons of the

road. In fact, it is possible that the contribution of the low-grade freight toward the fixed and general operating costs may make it possible to charge lower rates on the high-grade freight than would otherwise be imperative. This may be cold comfort to the shipper if his rate is in fact not reduced, but it should at least show him that he is not being injured by the low-grade business and that his complaint should be directed against his own high rate and not against the other low rate.

That taking on additional low-grade traffic at rates less than the prevailing scale is to the interest of the railroad with unused capacity is quite clear; a railroad operates at decreasing cost, and an increase in the volume of traffic, even of low-grade traffic, is greatly desired, for it means greater profits. As a matter of fact discriminations between goods started in this way. At first mostly passengers and high-grade freight were carried, but it was soon perceived that it was worth while to take the low-grade freight at a lower rate in order to spread the fixed and overhead costs over a larger volume of traffic. A railroad which has reached the point of maximum utilization does not have the same inducement to increase the volume of traffic. In fact it may find greater profit in raising the rates on the least profitable goods, thereby reducing the volume shipped, in order to have more carrying space for the more profitable freight, at least until such time as it appears wise to enlarge its facilities.

Railroad competition. It has been suggested earlier in this chapter that the railroads present the extreme case of an industry subject to the law of decreasing costs. In a previous chapter (Chapter XVII) the reader became acquainted with this type of industry and saw that the state of decreasing costs is an unstable one, in which there are no economic forces capable of fixing a definite price. Cutthroat competition is the result, continuing until equilibrium is attained through sufficient reduction of the number of competitors or through combination or monopoly. The railroads furnish the most instructive example of all this. Let us consider a hypothetical illustration.

The X & Y railroad runs between A and B, and has a monopoly

of the traffic. It carries, we will assume, 100,000,000 tons of freight at fifteen cents a ton, and has therefore a gross operating revenue of $15,000,000. Four million dollars cover the fixed charges, $4,000,000 the general operating expenses, and $4,000,000 (four cents per ton), the particular operating expenses, leaving $3,000,000 to be distributed to the stockholders in dividends. A new road, the M & N, is built between A and B to compete with the X & Y. It charges the same rate as the X & Y; *i.e.*, fifteen cents a ton, and naturally diverts some of the traffic from the X & Y — 25,000,000 tons, let us say. This leaves the X & Y in the following position:

Gross revenue		$11,250,000
Expenses		
Fixed charges	$4,000,000	
General operating expenses	4,000,000	
Particular operating expenses	3,000,000	
		11,000,000
Profits (available for dividends)		$250,000

The M & N railroad in the meantime is not satisfied with the share of traffic which it is getting and reduces its rate to twelve cents a ton, thereby attracting so much of the traffic of the X & Y railroad that the latter is forced to meet the reduction. Even if we assume that the X & Y regains its former share of the traffic its gross revenue will be only $9,000,000, while its expenses are still $11,000,000. Not only have profits been wiped out entirely, but only $2,000,000 is available for paying the fixed charges. The railroad is accepting traffic at a rate not sufficient to cover the total expenses of furnishing transportation, but it must continue to accept the traffic, since at least it yields something to apply to the fixed charges, and it is better to lose $2,000,000 a year than to lose the whole $4,000,000 of fixed charges, which would be the result if operations were suspended. And the road will be especially eager in seeking more traffic even at this unremunerative rate. If it could for example secure 10,000,000 tons additional freight, it would receive an income of $1,200,000 and incur an additional

expense of only $400,000, leaving $800,000 to be applied to the fixed charges.

The rate war may continue until the X & Y railroad has nothing at all with which to pay the bondholders, but this does not necessarily end the competition. The corporation is reorganized, the stockholders are wiped out, and the bondholders become stockholders. Inasmuch as the major portion of the fixed charges has been eliminated the road is in a better competitive condition than before, and the warfare continues until one side or the other decides that peace on any terms is to be preferred to continued struggle and allows itself to be absorbed by its stronger rival. Cessation of hostilities may come before this point is reached, on the basis of an agreement to maintain rates and pool either the traffic or the profits. In either case the result is the same — the elimination of competition.

That this competition may be disastrous to the railroads is self-evident; but it does not follow that the public is the gainer thereby. The shippers of goods at the competitive points doubtless have their markets widened for them by the lower rates, but this may prove to be harmful even to them in the long run, for a long drawn out period of cutthroat competition may lead to an artificial location of industries and their establishment on a scale which cannot be maintained under a rate system which is reasonably remunerative. The rise in the rates at the end of the rate war is certain to affect some industries adversely and may more than counterbalance the gain during the period of low rates. Society as a whole is bound to lose on account of the dislocation of industry caused by unremunerative railroad rates.

Rate discrimination. There will always be discrimination in the rates charged different classes of goods under the principle of charging what the traffic will bear, as we have seen. Two other kinds of discrimination are virtually sure to appear where there is competition between railroads.

Local discrimination. Such discrimination, in the broadest sense of the term, occurs whenever the rates between different points for a given class of goods are not proportional to the respec-

tive costs incurred. If the New York, New Haven & Hartford Railroad should charge the same rate for shipping a case of firearms from New Haven to Hartford, a distance of thirty-seven miles, as it charges from New Haven to New London, fifty-one miles, this would be a local discrimination. If it should charge more to Hartford than to New London, its action would be even more clearly discriminatory. Finally, if the rate to Springfield were equal to or less than the rate to Hartford, which must be passed through on the way to Springfield, we should have the extreme case, in which the shorter haul is included in the longer.

All variations from a flat rate per ton-mile, or other unit, are not evidences of local discrimination. The goods must be of the same kind (a different rate for coal and shoes is not local discrimination); they must be shipped in the same volume (it is quite proper to make allowance for carload shipments as compared with less than carload shipments); and the conditions of transportation must be essentially the same. We should expect the rate charged on trains running through mountainous country to be higher than rates charged on similar shipments through level stretches.

Examples. Local discrimination has been very marked in the past and still exists to no inconsiderable extent, and the situation sketched above is by no means without parallel in railroad history. "On first-class goods, at one time, the rate from Louisville to Montgomery was $1.26 per hundredweight. On to Mobile, 180 miles further south, it was only ninety cents. In the same territory the rate on kerosene oil from Cincinnati at times has been three times as much to interior points as to New Orleans, three times as far. . . . The rate on rope from San Francisco to Independence, Kansas, is seventy-five cents; while the same goods are hauled on through Independence, Kansas, much farther, to Missouri River points for sixty cents per hundred pounds. Wichita, Kansas, complains that cotton piece goods from New York by way of Galveston are rated at $1.36, while by the same route to Kansas City, 225 miles longer haul by that route, the charge is only ninety-three cents." [1]

[1] W. Z. Ripley, *Railroads, Rate and Regulation* (1912), p. 216.

The reason for local discrimination. The chief reason and the sole justification for local discrimination are found in the situation created by competition with another railroad or with a carrier by water. Two roads run between A and B, as shown on the diagram below.

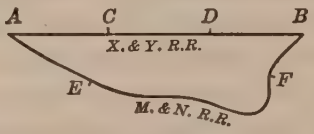

Fig. 1.

The X & Y railroad has a more direct route between A and B than does the M & N railroad. The former charges, let us say, three cents a ton-mile on freight from A to C, C to D, D to B, and also on the stretch from A to B. The M & N railroad likewise charges three cents a ton-mile on its intermediate traffic, but if it were to charge the same rate from A to B, the total charge per ton would be considerably greater than that from A to B via the X & Y. It therefore charges the same amount as the X & Y railroad even though this charge may be less than that from A to F. Shippers and consumers at F complain against this discrimination; they say, with truth, that more is charged for the short haul than for the long haul, even though the short haul is included within the long haul, and they further maintain that this practice is unjust and should be stopped.

The defense of the railroad rests on the assumption that some of the through traffic is essential, either to justify the existence of a railroad through the intermediate towns or to enable the intermediate rates to be as moderate as they are. The reasoning underlying this is essentially similar to that used to explain discrimination between goods. The rate on the through traffic is remunerative in the sense that it covers the particular operating expenses involved and makes a contribution toward the constant expenses. Without the through traffic this amount would have to come from the intermediate traffic alone, and the rates would therefore be higher than they now are, or in anticipation of such a situation the road would not have been built at all. The discrim-

ination, though real, is nevertheless essential to providing service for the intermediate points, and while the producers, shippers, and consumers at the intermediate points may be handicapped as compared with those at A and B, they are better off than they would be without the service or with still higher rates.

All discrimination between places cannot plead this excuse. Local discrimination is sometimes used to conceal personal discrimination in the interest of favored shippers, as will be noted later. Discriminatory rates to build up the land values of regions in which the railroad or officials own land have not been unknown; and for other reasons the railroad officials may arbitrarily favor one place over another.

Another example. An interesting case illustrating local discrimination is given by President Hadley in his *Railroad Transportation.* "On the coast of Delaware, a few years ago, there was a place which we shall call X, well suited for oyster-growing, but which sent very few oysters to market, because the railroad rates were so high as to leave no margin of profit. The local oyster-growers represented to the railroads that if the rates were brought down to one dollar per hundred pounds, the business would become profitable and the railroad could be sure of regular shipments at that price. The railroad men looked into the matter. They found that the price of oysters in the Philadelphia market was such that the local oystermen could pay $1 per hundred pounds to the railroad and still have a fair profit left. If the road tried to charge more, it would so cut down the profit as to leave men no inducement to enter the business. That is, those oysters would bear a rate of $1 per hundred, and no more. Further, the railroad men found that if they could get every day a carload, or nearly a carload, at this rate, it would more than cover the expense of hauling an extra car by quick train back and forth every day, with the incidental expenses of interest and repairs. So they put the car on, and were disappointed to find that the local oyster-growers could only furnish oysters enough to fill the car about half full. The expense to the road of running it half full was almost as great as of running it full; the income was reduced one half. They

could not make up by raising the rates, for these were as high as the traffic would bear. They could not increase their business much by lowering rates. The difficulty was not with the price charged, but with the capacity of the local business. It seemed as if this special service must be abandoned.

"One possibility suggested itself. At some distance beyond X, the terminus of this railroad, was another oyster-growing place, Y, which sent its oysters to market by another route. The supply at Y was very much greater than at X. The people at Y were paying a dollar a hundred to send their oysters to market. It would hardly cost twenty-five cents to send them from Y to X. If, then, the railroad from X to Philadelphia charged but seventy-five cents a hundred on oysters which came from Y, it could easily fill its car full. This was what they did. They then had half a carload of oysters grown at X, on which they charged a dollar, and half a carload from Y on which they charged seventy-five cents for exactly the same service.

"Of course there was a grand outcry at X. Their trade was discriminated against in the worst possible way — so they said — and they complained to the railroad. But the railroad men fell back on the logic of facts. The points were as follows: 1. A whole carload at seventy-five cents would not pay expenses of handling and moving. 2. At higher rates than seventy-five cents they could not get a whole carload, but only half a carload; and half a carload at a dollar rate (the highest charge the article would bear) would not pay expenses. Therefore, 3. On *any* uniform rate for everybody, the road must lose money, and 4. They would either be compelled to take the oyster car away altogether, or else get what they could at a dollar, and fill up at seventy-five cents. There was no escape from this reasoning; and the oyster men of X chose to pay the higher rate rather than lose the service altogether."[1]

Conclusion. When two railroads are in competition for the through traffic, there will almost certainly be local discrimination on at least one road, and there is very likely to be discrimination

[1] A. T. Hadley, *Railroad Transportation*, pp. 116–117.

against local points on both roads. In our example we assumed that no discrimination was found on the X & Y railroad, but as a matter of fact, this situation is scarcely to be expected, for this road will be tempted to try to regain some of the traffic lost to the M & N railroad, by lowering its through rates. In fact if the route by the X & Y railroad is very considerably shorter than the other, the former may be able to lower the rate to a point which will be absolutely unremunerative for the M & N railroad and retain all the through traffic.

Discrimination between places is a serious thing. Towns against which there is discrimination are injured, their industries may be destroyed, and their population driven elsewhere. The towns which enjoy the low competitive rates are given artificial advantages over the other towns which may offset or completely eliminate real natural advantages possessed by the latter. Population is attracted to the favored towns and the problems of city life intensified. Nor do the railroads themselves, as a whole, gain by the practice, for while a certain increase in traffic is likely to be stimulated by the lower rates, it is questionable whether it will be sufficient to compensate for the rate reduction. Nevertheless local discrimination is an inevitable feature of railroad competition, and it is probable that there would be a certain amount even under monopoly.

Personal discrimination. We have now to consider a third form of discrimination in railroad rates; namely, discrimination between persons. Such discrimination is peculiarly the result of railroad competition. Each road is forced to meet the cuts in the rates initiated by its competitors in order to get any business; if it can secure additional business by a secret contract with a particularly prominent shipper it is to its advantage, temporarily at least, to do so. Referring to our hypothetical example above, suppose that the X & Y railroad quotes a secret rate of ten cents a ton to a favored shipper at a time when the published rate is still twelve cents a ton and thereby diverts 10,000 tons of freight from the M & N. It will secure $1,000 of additional revenue at an expense of $400, and its rival will find its revenue reduced by $1,200 and its expenses reduced by only $400.

The essential feature of personal discrimination is to select some one shipper and attach him so firmly to the railroad by favors of one sort or another that he will patronize that railroad to the exclusion of all others. The tendency is quite naturally to favor the growth of industrial monopoly by restricting the area and scope of competition. Quite aside from its general social bearing, the growth of monopoly power among its shippers is not without danger to the railroad itself. When the fierceness of competition has died down and the railroad wants to readjust its rates, it may find itself no longer the master of the situation. The monopoly, ungrateful for past bounties, is in position to demand special rates by threatening to throw all of its business to the competing railroads.

Personal discriminations of all kinds — some of them straight rebates, and others concealed more or less successfully in allowances for demurrage, etc. — have been notorious in American railroad history. One of the outstanding examples is seen in the arrangement between the Standard Oil Company and the Cincinnati & Marietta Railroad. On shipments from Macksburg, Ohio, to Marietta the Standard Oil Company at one time not only paid but ten cents per barrel of crude oil, whereas the independent companies were charged thirty-five cents, but, more than that, the railroad turned over to the Standard Oil Company twenty-five cents out of each thirty-five cents paid by its competitors! In all fairness to the Standard Oil Companies of the present time it should be emphasized that this occurred many years ago, and that the Standard Oil Companies today are among the most scrupulous in their observance of the literal terms of the printed tariffs.

Local discrimination is sometimes used as a means of concealing personal discriminations. An influential manufacturer may represent the only important manufacturing interests in a small city; if open but discriminating rates are quoted on shipments from that city his competitive power is obviously strengthened. Of course it is possible for other manufacturers to settle in the city and secure the same rates, but the process is necessarily a slow one.

Other examples might be cited, but they are not necessary to demonstrate how totally at variance with the ideal of the public

utility is discrimination between shippers. It is quite contrary to all our notions of fair play that one man should be singled out and favored to such an extent that others are ruined and driven out of business, particularly by an agency upon which the whole public is dependent and which has received important privileges from the public. The encouragement to monopoly which results cannot be viewed with complacence, nor can it be urged, with any cogency, that the railroads act in accord with the public interest in simply hastening the process of natural selection among industrial leaders, for there is little evidence to show that efficiency is apt to be a consideration of any importance in selecting men or companies as the recipients of railway favor. That the public guards its own best interest in demanding that public authority put an end to these abuses is scarcely open to question.

The breakdown of competition. A picture of the many undoubted abuses and the many practices at least open to serious question which follow in the wake of unregulated railroad competition is now before us. It is perfectly clear that competition fails in the railroad business to do what it accomplishes on the whole in industry in general; namely, to ensure the production of the commodities and services which the people want at prices at least not far from cost of production. In the early days of railroading competition was relied upon to protect the public against the monopolistic practices of the railroads, and our early legislation attempted to compel the maintenance of competitive conditions. This childlike faith in competition was woefully misplaced.

Not only does competition fail to produce the results hoped of it while producing other results not foreseen or desired; competition proves itself either a pure figment of the imagination or at most an unstable and temporary situation, with an inherent tendency toward monopoly. Complete competition between railroads has probably never existed to any important degree; even under what would be called competitive conditions, a partial monopoly usually is found. If one examines the routes taken by the roads running between New York and Chicago he will realize that, while competitive rates exist on the through traffic, each of the roads has a more

or less complete monopoly of its local traffic. In fixing the rates on local traffic the zeal of a railroad to secure monopoly profits is tempered only by the realization of the fact that the producers of its region have to meet in the common markets the competition of producers of other regions, and that therefore they cannot charge rates so high as to make it impossible for the local shippers to meet this competition.

The most important reason for the failure of competition however is that competition is self-destructive; it degenerates into cut-throat competition and ends in combination or monopoly. As we have seen, the presence of so large a percentage of expenses incurred for all services makes it impossible to equate the service to the expense of rendering it, and the fact that so large a part of the expenses are constant in the face of a decrease or an increase in the volume of traffic places a premium on securing a large volume of traffic, up to the point of complete utilization of the plant. The result is that when keen competition starts the prime object of traffic managers is to secure traffic almost regardless of the rate. Rates will be cut, and they may fall so low as to cover only the particular operating expenses; secret rebates are given to attract shippers, while at the same time the shippers at the non-competitive points are charged monopoly prices. The end of this is the elimination of competition through consolidation or combination.

When once a road has gained a monopoly position, it is easy to hold. Even though the road may be securing monopoly profits which would ordinarily be more than ample to attract the investment of capital, it may be impossible to interest capitalists in the building of a competing railroad. Investors want to be very certain that the traffic is sufficiently large to assure both roads of an adequate rate of profit at the competitive rate on freight and passengers. The investment in a railroad is large, and when once made it cannot be transferred to other uses. Competition therefore gets under way very slowly and cautiously even in the presence of monopoly profits.

Monopoly inevitable and economical. There can be no escaping the conclusion that, on the whole and in the long run, monopoly

in the railroad business is inevitable. And the important corollary follows that railroad monopoly is more economical than competition. Not only does a lower unit cost result from the more complete utilization of existing facilities, but when additional facilities must be provided it is better to enlarge the existing plant than to build a new railroad. The laying of additional rails, the expansion of terminals, etc., entail, it is true, a considerable increase in the fixed charges, but the additional cost is much smaller than that necessary to develop the right of way, lay the rails, and provide the equipment necessary for a new railroad. In other words there are economies to be obtained from large-scale operation. A fully utilized two-track road operates at a lower unit cost than a similar one-track road, and the same holds true of a four-track road as compared with a two-track road. Monopoly allows us therefore to reap the benefits to be derived from complete utilization of the plant and to realize the economies which result from large-scale operation. It represents a more economical use of the limited supplies of labor and capital.

Regulated monopoly. While it is pretty clear that monopoly of rail facilities is the condition best adapted to the public interests from the standpoint of the most efficient utilization of labor and capital, unregulated monopoly is the cause, not merely of fears which may be somewhat irrational, but of misgivings which are well grounded on facts and principles. Unregulated monopoly carries with it no guarantee of adequate service at rates which are only reasonably remunerative or of impartial treatment of all shippers. In fact we may expect the railroads to consider profits before service, to be influenced by private motives rather than motives of public service, to forget that they have, by reason of their quasi-partnership with the public, obligations which should be fulfilled.

From a practical standpoint unregulated monopoly of the railroads is inconceivable in the United States. The choice lies, therefore, between (1) government ownership and operation and (2) government regulation of the activities of private companies. The problems arising out of this conclusion form the subject matter of the following chapter.

CHAPTER XXX

GOVERNMENT REGULATION OF RAILROADS IN THE UNITED STATES

The whole problem of transportation presents so large a subject that its complete treatment is quite beyond the limits of a general book such as this. Having obtained from preceding chapters a broad concept of the economic nature of railway transportation, we shall have to be content with a study in the present chapter of what has been accomplished in the United States alone in the field of government regulation of privately owned railroads.

State regulation of rates. The need for governmental regulation of the railroads was not clearly recognized in the pioneer days of railroad building and indeed did not find expression on the statute books until the years around 1870. In the decades preceding the primary need was felt to be the building of roads; competition could be relied upon to regulate them. Gradually however abuses sprang up which the public refused longer to tolerate, and legislation was passed in many states to control, to a greater or less extent, the activities of the railroads. It is significant of the spirit of the times that the individual states made the first move in this direction. The state of Massachusetts created its railroad commission in 1869, whereas no attempt was made by the federal government to assert effective control over the railroads until 1887, when it was virtually forced to enact legislation because of the states' lack of jurisdiction over interstate traffic.

The Massachusetts commission was of the supervisory type; it had no powers of compulsion but made investigations and recommendations to the railroads, relying upon the force of educated public opinion to enforce its decisions. Even with these

decidedly limited powers it attained a considerable degree of success, and it was of importance as a pioneer and an example for other states to follow. It was not however the prototype of the railway commissions of today, which are usually clothed with mandatory powers.

The movement for strict regulation. This movement found its origin in the agricultural states of the Middle West, which in the early seventies were suffering from an overexpansion of agriculture, as well as from the universal depression in business. The chief complaint against the railroads was that exorbitant rates were charged on shipments of agricultural produce, and when the competitive wars of the seventies somewhat reduced the force of this argument the charge of local discrimination took its place. There was moreover a profound conviction that the common attitude of the railroad executives, who appeared to regard themselves solely as the representatives of private interests, without obligation to the public and not amenable to the legislative power, was dangerous and needed curbing. The corruption evidenced in giving passes to those whose official position might enable them to lend support, in bribing legislatures, and in the skilful juggling of the finances of the roads for the benefit of insiders added to the force of the movement.

The "Granger laws." Early in the seventies many of the states of the Middle West passed the Granger laws (so called because of the support which the National Grange of the Patrons of Husbandry lent to the movement) forbidding discriminations, and prescribing maximum rates which might be charged. Naturally enough the railroads at once contested the constitutionality of laws which threatened to curb their arbitrary powers. We cannot enter into all the details of these cases, but we may note several points of fundamental importance which were established in the final decisions.

The first question requiring adjudication was whether the state possessed the power to regulate the roads to the extent of fixing maximum limits beyond which rates might not be raised. The railroads based their cases on the argument that the fourteenth

amendment to the federal constitution, which forbids any state to deprive any person of property without due process of law, rendered rate laws unconstitutional. The Supreme Court in its decision noted that it had long been customary both in England and in the colonies for legislative authority to prescribe maximum charges for ferries, common carriers, and the like. It was not impressed with the argument that fixed maximum rates deprived persons of property without due process of law, and it further laid down the doctrine that the railroads as common carriers were affected with a public interest which it was the inherent right of the state to safeguard.

The second contention of the railroads was that the state, granted it had the authority to fix rates, did not have final authority in determining reasonable rates. The railroads wanted the right of appeal to the courts for a review of specific rates which they might consider unjust. This the courts denied, affirming that the legislature had supreme and sole authority to determine the reasonableness of rates and that the courts had no jurisdiction. In the course of time, however, this doctrine was overthrown, and the right of the courts to review rates is now accepted as fundamental and applies to all rates set by federal and state legislatures and commissions.

Many of the Granger laws were ephemeral. Conceived in a spirit of bitterness and even of hostility to the railroads, it cannot be questioned that they overshot the mark and reacted adversely on the interests of the railroads and of the public. Their chief services lay in bringing about an enunciation of first principles, in blazing a path in the unknown and untried field of railway legislation, and in opening the eyes of the railroad executives to the possibilities of the future.

The conflict of jurisdiction. Conflict between state and federal authorities over the control of the railroads was inevitable, so soon as the railroads emerged from that condition where they consisted of short disconnected lines only and became increasingly important as the carriers of commerce between the several states. The United States constitution gives to the federal government

control over foreign and interstate commerce, while each individual state has authority over the commerce originating and having its entire course and destination within its boundaries. The difficulty of fixing rates for intrastate trade without such action having an influence on the interstate rates, or the reverse, is quite apparent.

In the Granger cases it was decided that, in the absence of federal legislation on the subject, the state had the power to prescribe rates to regulate intrastate traffic, even though such action might indirectly affect interstate commerce. But the Wabash case, decided in 1886, established the principle that, even though the federal government did not choose to exercise its power, it nevertheless possessed sole authority to regulate interstate commerce. Since it was inconceivable that interstate commerce should be left to the arbitrary control of the railroads, particularly since about three quarters of the traffic of the roads in 1886 was between the several states, Congress was forced to enact its first comprehensive regulatory legislation — the Act to Regulate Commerce of 1887.

With two authorities determining rates within a state — the federal authorities in control of interstate rates and the state authorities in control of intrastate rates — the question of the dominant authority had to be settled. In case of conflict should intrastate or interstate rates rule? In the Shreveport Rate Case (1914) the complaint was made that the rates prescribed by the Texas Railway Commission on intrastate traffic constituted undue discrimination against traffic coming into the state from Shreveport, Louisiana. The court, in its review of the case, sustained the findings of the Interstate Commerce Commission that the state rates were unduly prejudicial to interstate commerce originating at Shreveport. It also sustained the Commission's fixation of reasonable maximum interstate rates. Based on these two sets of items, the court required the railroads to eliminate the undue prejudice found to exist. This they were at liberty to do by lowering the interstate rates to the level of the state rates. They were also expressly permitted to eliminate the undue prejudice by observing on both classes of traffic the interstate rates set by the Commission, even though this involved the disregard of the

Texas state rates. The Shreveport case marks therefore a step towards more complete control over rates by federal authority The final step in this direction, short of complete control over intrastate rates, was taken in the Transportation Act of 1920, which states that where a state-made rate discriminates against interstate traffic, the Interstate Commerce Commission shall prescribe a rate to remove this discrimination, which shall be binding, " the law of any State or the decision or order of any State authority to the contrary notwithstanding."

The growth of federal regulation has meant a decrease in the scope of state regulation, but few would question the suitableness of this. The multiplicity and diversity of regulations which might apply to the railroads were the states permitted to attempt to foster home industries by such means would create an intolerable situation. There must be some mediating force, some authority able to reconcile the differences between the states. The supremacy of the national government in the field of railroad legislation does not mean however that it is useless for states to attempt any measures of control. State railroad boards are still useful for purposes of local control, of gathering information, and of representing in an official way the interests of the state. Doubtless they will in the future be more engrossed in the task of controlling public utilities which are purely local in character, but that they will cease to interest themselves in the railroads is neither necessary nor probable.

Federal regulation: the Interstate Commerce Act of 1887. It will be sufficient for our purposes to summarize the most important of the provisions of the Act to Regulate Commerce of 1887. The act was made applicable to interstate and foreign freight and passenger traffic by rail or by rail and water; intrastate commerce was specifically excluded. Rates were to be just and reasonable; and personal, local, or commodity discriminations were declared unlawful. Each road must provide reasonable and equal facilities for the interchange of traffic with other lines, and it might not discriminate in rates between such connecting lines.

Charging, under substantially similar circumstances and condi-

tions, more for the transportation of passengers or like kind of property for a shorter than for a longer distance over the same line in the same direction, the shorter being included within the longer distance, was declared unlawful. It was however provided that the Interstate Commerce Commission, created by this same act, might after investigation afford the carriers relief from this provision if special conditions seemed to warrant it.

An attempt was made to ensure the continuance of competitive conditions by making unlawful any contracts or agreements between railroads to pool freights or to divide earnings.

Rates and fares were to be printed and posted in every station and filed with the Interstate Commerce Commission. Rates not so posted were unlawful. No advance in rates might be made without ten days' notice, but a reduction might be made at any time, provided immediate notification were given the public.

Damages might be recovered from a railroad by a person injured through unlawful practices, and provision was made for a fine of not more than $5,000 for each offense upon conviction in the courts. Two years later imprisonment was made an alternative or an additional penalty for certain offenses.

Finally the Interstate Commerce Commission was created as a supervisory and regulatory body. Under the law as amended to the present, the Commission consists of eleven members appointed by the President by and with the advice and consent of the Senate, for terms of seven years. Each commissioner receives a salary of $12,000 a year. He is forbidden to engage in any other activity whatever, and, to ensure absolute impartiality and to protect the members from criticism, no commissioner may own stocks or bonds in any railroad company.

The Commission was given by this act the right of investigation and report, either on its own motion or on the complaint of a shipper. It could compel witnesses to testify and to produce books and papers, it could order unlawful practices to cease, and if necessary it could apply to the federal courts to issue an enforcing order to the offending road.

Defects of the Interstate Commerce Act. The Interstate Commerce Act did not work out altogether as anticipated by its framers. Many defects in the law itself were discovered, and the interpretations given by the courts to some of the clauses seriously restricted the powers of the Commission. The first test of the authority of the Commission arose over the question of its power to compel reluctant witnesses to testify. Adverse decisions by the courts forced the passage of additional legislation in 1893, but it was not until 1896 that the right of the Commission to compel testimony and the producing of evidence was fully established by a decision of the Supreme Court.

Another grave defect was that the Commission's orders were not binding upon the carriers. They might obey them or not as they chose, and the only recourse was to take the case to the courts. If the courts were satisfied that a lawful order had been disobeyed, they might issue an injunction enforcing the order. But the delay involved in instituting such suits and in carrying an appeal to a higher court, as well as the fact that the courts allowed the railroads to submit evidence which had not been previously submitted to the Commission in its own investigation, made this method of regulation highly unsatisfactory.

The final blow came in 1897 when it was decided that the Interstate Commerce Act did not give the Commission the right to determine maximum rates. Henceforth it could, subject to review of its action by the courts, determine whether a particular rate was reasonable or not, but if it condemned a rate as unreasonable it could not fix the maximum that might be charged. A railroad could, by a series of minute rate shadings, postpone almost indefinitely the reduction of a rate to that point considered by the Commission to be just.

Development of regulation to the period of war control. The action of the courts in thus limiting the powers of the Commission and a widespread feeling that more rigorous control of the railroads was desirable led to the passage of supplementary legislation, and beginning in 1903 the Act to Regulate Commerce of 1887 has been frequently amended, either to attain the objects of the act or to

bring the railroads under greater control than originally contemplated.

As we have seen, the rate-making power of the Commission was for all practical purposes abolished in 1897. Shippers could still appeal to the courts against unreasonable charges, but this was a right which in the nature of the case would seldom be invoked. To remedy this situation the power to determine and prescribe maximum rates in the place of existing rates which it considered unreasonable was explicitly given the Commission. It might moreover prescribe maximum joint rates on through routes and settle the division of the rates between the several carriers concerned.

The railroads' power to alter rates was circumscribed, first by requiring thirty days' notification of any alteration in the published rates and later by giving the Commission authority to suspend the operation of any new rate, classification, regulation, or practice for a period of 120 days, with an additional period if necessary, while it made an investigation of the need or justice of the proposed change.

Equally important were the improvements in the procedure devised to carry out the orders of the Commission. In order to decrease the time required to settle cases the Expediting Act was passed, giving precedence in the Circuit Courts to suits in equity arising under the Interstate Commerce Act or the Anti-Trust Acts in which the United States appeared as complainant or where the Attorney-General certified that the public interest was involved. Appeals from the decisions of the Circuit Court were to be taken directly to the Supreme Court within sixty days after the decree had been given by the lower court.

Orders of the Commission, except for the payment of money, were made effective within thirty days after the date of issue and remained in force for two years unless suspended by the Commission or the courts. The railroads were debarred from pursuing their former tactics of obstruction by the imposition of a fine for neglect to obey an order of the Commission amounting to $5,000 for each offense or each day of the same offense. In case of failure

to obey the orders of the Commission the Circuit Courts were furthermore directed to enforce compliance.

Protection was afforded the carrier against arbitrary or unreasonable action of the Commission; it had the right to request the court to suspend an order of the Commission and to appeal the decision of the Circuit Court directly to the Supreme Court. In no case however was an order of the Commission to be suspended except on hearing, after not less than five days' notice had been given the Commission. Since 1906 the courts have contented themselves with reviewing questions of law and have considered the Commission as a competent and effective administrative board, whose findings of facts and whose orders should not be questioned so long as they are confined within its lawful field. This has ended the troublesome question of admitting fresh evidence in the court hearing, a practice which not only prolonged the settlement but made conflict of decisions between the Commission and the courts a frequent occurrence.

The provisions intended to eliminate personal discriminations were strengthened, in the first place, by making the published tariff the test of lawfulness, so that a rate other than the published tariff was *ipso facto* illegal, regardless of whether other rates had been charged or not. Previously, in order to prove the existence of discrimination, the Commission had found it necessary to show that a secret and preferential rate had been given in some specific case and that higher charges had actually been levied on like and contemporaneous shipments. In the second place, the penalties for the offer, solicitation, or acceptance of rebates were made much more stringent than formerly. The Commission was also given greater authority over long and short haul charges in order to prevent unjustifiable discriminations.

Maintenance of competition. The maintenance of competitive conditions on the railroads has been a cardinal feature of American policy. We have already noted the legislation against pooling. Pools were formed at quite an early date to avoid the ruinous effects of cutthroat competition, and some of them were quite successful in spite of the fact that they had no standing before the

law. According to the common law doctrine, agreements or contracts in restraint of trade — and pools to divide earnings or freights were open to this interpretation — are null and void and cannot be enforced by court action. Declared illegal and made punishable under the federal law, pooling gave way to traffic associations designed to control and maintain rates and to combinations or consolidations of competing lines. To meet this situation the government invoked the authority of the Sherman Anti-Trust Act of 1890 which declared illegal "every contract, combination in the form of trust or otherwise, or conspiracy in restraint of trade or commerce among the several states, or with foreign nations," and in a number of notable cases succeeded in forcing the dissolution of the combination.

In recent years however the fear of a railroad monopoly has somewhat died down. There is perhaps less to be gained from competition between railroads, as the decrease in the unused capacity of the roads has in large part removed the incentive to cut rates in order to secure traffic. Certainly the more effective control exercised by the Interstate Commerce Commission has contributed to this change of attitude, as well as a somewhat belated appreciation of the inherent difficulties of maintaining competitive conditions and of the desirability of securing the advantages which may come from unified control and the elimination of needless duplication in facilities and services.

Limitation of rate increases. One effect of the regulation of rates, unforeseen when the legislation was passed, worked great injury to the railroads. Most of the legislation came during the latter years of the nineteenth century, when the general price level was, as we have seen,[1] at its lowest. Since then there has been an extraordinary increase in prices in general. Railway expenses rose with all prices. Supplies and equipment used by the railroads increased in price, and the higher costs of living made it necessary to grant wage increases. Railway income did not increase proportionately because of the legal restrictions upon rates. However

[1] *Cf.* Chapter XXIV for an account of the changes in the general level of prices since the Civil War.

distressing this situation was to the railroad managers, it was generally viewed with equanimity by the public and by the Interstate Commerce Commission, which refused to sanction rate increases sufficient to cover the extra costs. The result was little short of disastrous. Railway profits declined and were in many cases succeeded by losses. Investors were reluctant to put their capital into railroad securities, their timidity being increased, unfortunately, as a result of serious financial mismanagement of some of the more prominent roads. The consequence was an impairment of credit which made it impossible to procure funds for needed improvements, extensions, and equipment.

War time operation by the government: the railway situation. During the World War and for a number of months thereafter the government assumed entire control of the operation of the steam railroads and the services ancillary to their operation, such as the express business, sleeping cars and parlor cars, telephone and telegraph lines, etc. The reason for this step was the manifest inability of the railroads to handle the enormous volume of traffic with the dispatch necessary in war times.

In the years before the outbreak of the war, as we have seen, the costs of operation had been mounting without a corresponding increase in traffic or revenue, and with disastrous consequences to the railroads. With impaired facilities and strained credit the railroads of the country were called upon to meet the transportation needs suddenly created by the World War. The roads were not only required to handle an increasing volume of export traffic, but the withdrawal of coastwise shipping to the transatlantic service placed upon them the burden of carrying an additional amount of domestic traffic. The relatively large amount of east-bound traffic as compared with the westbound and later the reduction in shipping caused by the submarine campaign brought about serious congestion in the terminals on the eastern seaboard and a car famine in the West. In addition the efficiency of the roads was decreased by an abnormally high turnover of the labor force. Wages in industrial lines were higher than on the railroads, and many skilled workers were diverted to munition plants and other industries.

The entry of the United States into the war intensified the problems. To the demand of the Allies for transportation services was added the demand of the various branches of the United States government for the transportation of soldiers and enormous quantities of freight for the building of military camps, for use in shipyards, etc. The railroads voluntarily formed a committee, called the Special Committee on National Defence of the American Railway Association, to bring about efficiency in service and coördination of effort, but, in spite of their achievement in increasing the tonnage handled, it became so apparent toward the end of 1917 that the transportation system was in imminent danger of a breakdown that the Interstate Commerce Commission sent a special message to Congress advocating the unification of the railroads, either by the carriers themselves with the coöperation of the government or by their operation as a unit by the President.

Government operation. On December 26, 1917, President Wilson issued a proclamation to the effect that he would assume control of the steam railroads two days later. Mr. William G. McAdoo, the Secretary of the Treasury, was appointed Director-General of Railroads and entrusted with the gigantic task of organization and operation.

While President Wilson was enabled to assume control of the roads by virtue of war powers which had been conferred on him by the act of August 29, 1916, the operation of the roads was carried on under the Federal Control Act, passed on March 21, 1918. This act provided for compensation to the railroad owners for the period of government operation, and for maintenance, repairs, and renewals; it empowered the President to initiate all charges, classifications, regulations, and practices and deprived the Interstate Commerce Commission of its power to suspend changes of rates; it created a revolving fund of $500,000,000 to pay the expenses of federal control and provided, finally, that the roads should revert to their owners twenty-one months after the proclamation by the President of the mutual ratification of treaties of peace.

Under federal operation the railroads were unified and their physical resources — terminals, equipment, and repair shops —

were pooled. In pursuance of the understanding that its duty was not only to expedite traffic but also to cut down the volume of traffic, the Railroad Administration reduced the passenger service where possible, increased demurrage charges in order to secure prompt unloading of freight cars, arbitrarily routed and distributed shipments, required permits for shipments, issued priority orders, etc. It effected economies in operation through consolidated ticket offices, centralized purchasing, and standardization of equipment. To handle labor, which was both discontented and inadequate to the needs, it set up a Board of Railroad Wages and Working Conditions, extended the eight-hour day throughout the railroad service, raised wages and standardized them throughout the United States. The increased cost of operation was offset in part by increases in freight and passenger rates; both interstate and intrastate rates were advanced.

The essential purpose of the federal administration was to keep men and materials moving smoothly and rapidly to their destinations. This was accomplished successfully; and in appraising the results of the period of federal operation it is essential to keep this point in mind, for from the standpoint of financial operations the Railroad Administration was not successful. Huge deficits were piled up, approximating $1,200,000,000 for the twenty-six months of federal operation.

The return to private operation. The financial tangle and the natural reaction to war time control with its many inconveniences, as well as a conviction on the part of many that government operation in itself is undesirable, led to the return of the railroads to private management early in 1920. In May, 1919, President Wilson announced that the railroads would be returned to their owners on January 1, 1920. The necessity of finding a more acceptable solution to the railroad problem than had hitherto been attained and the widespread interest of the general public in the question led to a careful survey of the possible conditions under which the roads should be operated in the future. The possibility of continuing government operation was not immediately discarded. In fact a certain portion of the people wanted

the ownership of the railroads, as well as their operation, vested in the government. Railroad labor, which had found the government a more sympathetic master than the private companies, was especially insistent on this and to this end worked out a fairly detailed plan, commonly known as the Plumb Plan. But the advocates of this plan were not numerous and influential enough to win Congress to further experimentation in government operation.

The Transportation Act of 1920. This law, known also as the Esch-Cummins Act, was passed as an amendment to the Act to Regulate Commerce of 1887. It represents an honest and sincere effort on the part of its framers to provide a solution of the railway problem which should be just to the owners of the railways and satisfactory to the public. In many respects it departs from the standards set in previous legislation, particularly in matters of policy. A peculiar importance attaches to it and to its success or failure, for the inability of the railroads to provide satisfactory service and to maintain such relations with their laboring force as to ensure continuity of service will unquestionably add to the force of the movement to nationalize the railroads. We shall therefore examine the provisions of this act in some detail.

Provision for return of the roads. The first matter to be adjusted was the actual return of the railroads to their owners, and it was provided that this should take place on March 1, 1920. In order to give the railroad managers time to adjust themselves to the changes which had come about under federal operation, the government offered to guarantee for six months the same rate of return which it had paid while in control of the roads. If any road which took advantage of this offer should secure earnings in excess of this the surplus was to revert to the government. A fund of $300,000,000 was appropriated to be loaned to the railroads for the purpose of making necessary extensions to their own equipment and of meeting maturing bonded obligations, and machinery was devised for the mutual settlement of the claims of the railroads and the government.

Financial provisions. It was clearly recognized that rates must be adjusted so as to give the railroads revenue adequate to attract

investment in their issues of securities. At the same time it was realized that a rate schedule which would assure this income to the majority of roads would yield to some a revenue far in excess of their needs. The problem was solved as follows: The Commission was empowered to divide the railroads of the country into groups and to establish rates which would allow each group as a whole to earn from five and one half to six per cent on the value of their property used in transportation. The law thus determined what it considered a fair return on the property; but after March 1, 1922, the Commission was given broad powers in determining a fair return, taking into consideration the transportation needs of the country and the necessity of expanding railroad facilities. At present it has fixed upon five and three quarters per cent on the value of the property as constituting a fair return.

It is worth emphasizing that these provisions do not in any sense guarantee a fixed rate of return to a given railroad. A strong railroad might earn from ten to twenty per cent on the value of the property, and a weak railroad might have no net earnings or even show a loss. It is expected however that the average earnings will be as high as five and three quarters per cent.

In order to guard against excessive earnings the so-called "recapture clause" was inserted. A railroad which earns more than six per cent on the value of railroad property used in transportation is required to turn over one half of the excess to the government. The railroad's share must be placed in a reserve fund, and it cannot be used save for the payment of dividends, interest, or rent of leased roads, until the fund equals five per cent of the value of the property. The government's share goes to a revolving fund to be used either (1) for loans to needy railroads to meet expenditures for capital account or (2) to refund maturing securities originally issued for capital account or to purchase equipment to lease to the railroads.

Several other provisions of this act amplify the control of the Commission over rates in other respects than implied above. The Commission is specifically given power to prescribe minimum rates. The importance of this provision lies in the recognition

of the desirability of calling a halt to cutthroat competition. The Commission may therefore refuse to allow a railroad to set a rate which would be manifestly unfair to a competing road. Mention has already been made of its power to establish intrastate rates when the existing state rates unjustly discriminate against interstate or foreign commerce. And lastly the Commission's power to relieve a carrier from the rigid interpretation of the long and short haul rule was somewhat circumscribed by the proviso that the Commission shall not allow the establishment of any rate which is not reasonably compensatory for the service rendered.

Labor provisions. Special provision is made in the Transportation Act for settling railway labor disputes, but it is not the first time that Congress has legislated on this subject. In 1888 Congress provided for voluntary arbitration in interstate railway labor disputes, relying upon the force of public opinion, informed on the subject by the report of an investigating committee, to impel acceptance of the award. In 1898 the Erdman Act made it the duty of the Chairman of the Interstate Commerce Commission and the United States Commissioner of Labor, upon application of either party to a controversy, to offer their services to bring about an amicable adjustment of the trouble through mediation and conciliation. Failing to do this, they were to try to bring about arbitration proceedings. Up to 1913, 61 cases were settled under this law, 26 of these through mediation, 10 by mediation and arbitration, and 6 by arbitration alone. In 1913 the Newlands Act amended and amplified the Erdman Act by providing among other things that the board of arbitration should limit its decision strictly to the issues involved. Between 1913 and 1917, 71 cases were settled under this law, 52 of them by mediation and 6 by mediation and arbitration.

On January 1, 1917, the Adamson Act establishing the eight-hour day went into effect and should to all intents and purposes be regarded as an adjustment of a labor dispute between employers and employees. During the period of war control a number of commissions and boards were set up to study wage conditions and to settle disputes as to wages and working conditions. These did

not involve special legislation, but they are of significance in that they are symptomatic of the attitude and power of organized labor which made it imperative to incorporate provisions for the settlement of labor disputes in the Transportation Act of 1920.

Two types of boards for the settlement of labor disputes are provided for in the Act of 1920. The first type, the Railroad Boards of Labor Adjustment, may be formed by voluntary agreement between the carrier and its employees. If established, the Board is required, at the request of the carrier, the employees, or the Railroad Labor Board, to hold a hearing, take testimony, and render a decision on disputes involving only grievances, rules, or working conditions. These Boards have the right to summon witnesses, compel testimony, and require the production of documents; they have no authority to enforce their decisions or to settle wage questions.

The establishment of the Railroad Labor Board was required. It consists of nine men appointed by the President for terms of five years; three men to be selected from the nominees of the carriers, three selected from the nominees of the employees, and three appointed directly by the President as representatives of the public interest. The Board has jurisdiction over (1) disputes concerning grievances, rules, or working conditions where the local Adjustment Board has been unable to render a decision, (2) disputes concerning grievances, rules, or working conditions arising where there is no local board, or where the Railroad Labor Board thinks it advisable to intervene, and (3) all disputes concerning wages. A majority of the Board must concur in any decision, and in wage awards at least one member representing the public must concur. The Board is given authority to compel testimony and the producing of evidence, but it has no power to compel the acceptance of any of its awards, other than that which it creates itself by informing the public of the merits of the points at issue and by rendering decisions which commend themselves to public opinion.

These provisions have not perhaps been long enough operative to permit a just appraisal of their effectiveness, but it may be said

that they have not worked out altogether as anticipated. The formation of the local boards has been very slow, largely because of the inability of the carriers and the railroad employees to reach an agreement as to the basis of their organization, and this has made it necessary for the Railroad Labor Board to consider a multitude of problems which otherwise might have been settled without reference to it. Furthermore the constitution of the Board, while admirably calculated in theory to preserve a just balance between the various interests involved, is open to criticism. The Interstate Commerce Commission has earned an enviable reputation for impartiality and ability to view in a broad and understanding way questions presented to it. Unquestionably this is partially due to the non-partisan character of the Commission. Unfortunately members of the Labor Board have felt very keenly their position as interested parties, and this has not been conducive to amicable and speedy settlements. It has been suggested that a separate Labor Board is in any circumstances a mistake. Questions of rates and wages logically go together. Should not the wage disputes be referred to the Interstate Commerce Commission, or to a subdivision of it created for this purpose?

Railway consolidation. In an earlier section mention was made of the change in public policy with respect to the maintenance of competition between railroads. This is nowhere more clearly illustrated than in the section dealing with consolidations. The Interstate Commerce Commission is directed to draw up a plan for the consolidation of the railroads into a number of large systems. It is intended to preserve competitive conditions and at the same time so to arrange the systems that the same rates on traffic will provide substantially the same rate of return upon the values of the respective properties. Competition will therefore be in service rather than in rates. The success of this plan will depend upon the attitude of the railroads, for the Commission has no authority to force consolidation, and it is not at all impossible that the financially powerful roads may be skeptical of the advantages to be derived from absorbing one or more struggling roads.

The railroads themselves may initiate steps toward acquiring

control of other roads. Consolidation or control through lease or stock ownership may be sanctioned by the Commission, in spite of state regulations or laws to the contrary. Furthermore pooling arrangements are allowed with the approval of the Commission, if they do not unduly restrict competition and if a gain to the public can be shown. Interlocking directorates are forbidden, but again broad latitude is given the Commission in the administration of this provision.

Railway securities. The exclusive right of controlling new issues of rail securities is given to the Commission. State control had proved itself in many cases to be arbitrary and discriminatory, and control by the federal authorities was adopted to ensure uniformity between the different railroads and to prevent the financial manipulation which was possible under the lax laws of some states. Power is given to determine whether the purpose to which it is proposed to apply the proceeds of the issue is "reasonable, necessary, and appropriate" for the proper performance of public service, and if the application is approved the Commission may make such terms or conditions governing the issue as it deems wise.

Railway service. Exceedingly broad powers with respect to car service were conferred on the Commission. It may regulate the movement of cars and locomotives, compel the purchase of all classes of equipment, determine the compensation which one road shall pay for the use of the equipment of another, compel the common use of cars, locomotives, and terminals, and in cases of emergency issue priority orders or embargoes on traffic. The extension or abandonment of service facilities is likewise under its control; no new railroad may be built, no existing road may be extended, nor may any line be abandoned without the consent of the Commission.

Regulation on trial. It may safely be said that our past experience has taught the futility of railway competition and the impracticability of leaving the transportation business in private hands without regulation. The only alternatives before us are (1) private ownership under government regulation and (2) government own-

ership. There is always present a considerable sentiment in favor of government ownership, though the American public as a whole is by no means won over to this policy. Rather we are committed to a further trial of government regulation. Past blunders in regulation have taught us some lessons. The Interstate Commerce Commission is clothed with almost unlimited powers of regulation. Indeed so extensive are these powers that some are apprehensive lest all initiative be taken from the railroad management. If failure attend railroad management in the coming years, either with regard to service or finances, it cannot be ascribed to the lack of regulatory authority. It appears that we are now likely to have a real test of the efficacy of government regulation.

CHAPTER XXXI

INDUSTRIAL MONOPOLY AND ITS CONTROL

In Part I of this book (Chapter VII) we obtained a view of the extraordinary development of large scale production and industrial combination and sought to appraise the economic advantages accruing therefrom. We were able to envisage the great economies which, within limits, large scale production offers. As to combination, we concluded that, while there is undoubtedly some opportunity here for saving in cost of production, the true economy from mere combination of separate business units is on the whole neither universal nor very great. On the contrary it appeared that the motive for combination and the explanation of its most conspicuous successes was generally possession of monopoly power. At another place (Chapter XVIII) we have studied the forces which govern the prices charged by monopolies. Our attention in the present chapter will be devoted to a more searching analysis of the nature of monopoly and an investigation of the broad economic aspects of industrial monopoly.

Definition and classification. *Monopoly*, as the term is used in economics, is the exclusive control of the supply of an economic good. A variety of examples was presented in Chapter XVIII, and others will occur readily to the reader. The term is also used (without ambiguity) as a synonym for *monopolist; i.e.*, the one who possesses such exclusive control of supply. What gives special importance to monopoly is the fact that it endows its possessor with the power to control price.

As an aid to our analysis, the following classification of monopolies will be useful:

1. Legal monopoly
   a. Public legal monopoly
   b. Private legal monopoly

2. Natural monopoly
   a. Possession of the sole natural source of a good
   b. Possession of a unique non-reproducible good
   c. Personal monopoly
   d. Capitalistic monopoly

Legal monopolies exist by virtue of rights or powers granted by authority of government, while natural monopolies grow out of the operation of natural forces, frequently in the face of governmental opposition and public disapproval.

Public legal monopolies. The government may reserve for itself the right to sell some commodity or service. In the United States the postal service of the first class — the carrying of letters — is exclusively in the control of the federal government; in France the tobacco business is a monopoly of the national government, and in many countries the sale of salt has been made an exclusive privilege of the government. Such monopolies we call *public legal monopolies.* The motive underlying the creation of a monopoly of this type may be purely fiscal; that is, the government may want monopoly profits in order to add to its revenues. But in many cases, indeed in most cases in modern times, the purpose is to safeguard the public health or in some other way to further the general welfare. Thus a government monopoly of the sale of liquor generally represents an attempt to find a more acceptable solution of the liquor problem than has been found under private enterprise. The municipality which operates its own water system has usually been led to take this step in order to be sure that the residents get pure water at a fair price.

Obviously public legal monopolies do not raise fundamental questions of regulation. If monopoly profits are earned, the presumption is that the people benefit by lower taxes. The fundamental problem here is whether it is to the interest of society to have the government engage in a given branch of industry. We shall give further attention to this problem in connection with our study of government industry (Chapter L).

Private legal monopolies. On the other hand exclusive privileges are frequently conferred by a government on some person or

corporation. In former days private legal monopolies were often bestowed by a sovereign on some courtier as a mark of royal favor, but in modern democratic countries there is always at least the presumption that the special privilege is granted in the public interest and that the private gain is only incidental. In the United States the chief examples of private legal monopolies are patents, copyrights, and exclusive franchises.

Patents. A person who has devised a new machine or process may secure from the United States Patent Office a patent, which will give him the exclusive right to manufacture and sell the invention for a period of seventeen years, and he may further protect his invention by securing patent rights in other countries as well. During the period covered by the patent he may make such use of the right as he chooses, even to the extent of transferring his right to another, but when the patent right expires anyone may engage in the manufacture and sale of the device.

The patent monopoly is based on the theory that society as a whole is the chief beneficiary of all improvements in the instruments of production and that the chance of exclusive private gain which the patent imparts is necessary to induce men to undergo the labor of invention and the risks which are involved in experimentation and in manufacturing and marketing a new device or process. The gain to society, it is believed, more than outweighs the monopoly gain which accrues to the inventor. And when the period of the patent has expired, the whole benefit of the new device reverts in perpetuity to the public.

These fundamental principles upon which the patent rests are perfectly clear, and their soundness is scarcely open to question. Nevertheless the patent monopoly does present certain special problems. Patent rights have sometimes enabled monopolies to be founded which were not contemplated by the framers of the patent laws. A combination which controls the essential patents in an industry may acquire a position which cannot successfully be challenged even when the patent rights have expired. Again control of patented machines may enable a concern to secure a monopoly in unpatented machines as well. A company making a

patented machine and leasing it to manufactures may stipulate that another machine on which the patents have expired be used with it even though there be available equally good machines of other makes. Though such "tying" contracts are now unlawful in America, they were once fairly common. Again it has not been uncommon for large concerns to buy up new patents and suppress them, in order to continue reaping profits from an inferior device or to prevent competition from users of the new device.

Nor can it be argued that inventive genius will always be adequately rewarded. An inventor usually needs outside capital for the commercial development of his idea, but capitalists are loath to start competition with a strongly entrenched concern unless the invention is of unusual promise. The only market for the invention may be that offered by the concern already in control of the field; it is inevitably so if the invention is an improvement of a machine of which this concern holds the basic patent rights. It follows that the purchase price will be a non-competitive price, and if the company is interested in buying the patent only to suppress it, the royalties from the manufacture of the machine may be nonexistent or long delayed. Competent observers seem agreed that our patent laws need revision in order to give a more effective stimulus to invention and at the same time to prevent the growth of unintended monopolies based upon them.

The copyright. Monopoly of the publication and sale of books and other forms of literature and art is conferred by copyright. In the United States the monopoly runs for a period of twenty-eight years, with the privilege of renewal for another twenty-eight years. Like the patent, it is based on the principle that the production of worth-while books is thereby stimulated and that the public benefits in consequence.

The exclusive franchise. In order to secure the economies in operation which come from unified control, a franchise is sometimes granted by a city to a private company giving it the exclusive right to use the public streets to lay rails for a street railway line, or to erect telephone lines, or to lay water mains, gas pipes, etc. These

public utilities are in fact generally natural monopolies, and the exclusive franchise is but a recognition of the economic desirability and the practical inevitability of monopoly. If conferred at the inception of the service, exclusive franchise has the advantage of sparing society the wastes of cutthroat competition between rival companies.

The problem of control of such local utilities, as in the case of the railroads, centers around the regulation of rates and the quality of service, the authority being lodged either in the town or municipal government or in a public utilities commission created by the state. No novel questions of principle are involved in the control of the local public utilities.

Natural monopolies. We have found it convenient to classify natural monopolies in four groups. One form rests on the possession of the only natural source or sources of supply of some commodity or service. The owner of a mineral spring the waters of which have unique medicinal properties has a monopoly of this type. A railroad company which owns the only practicable pass across a mountain range has a natural monopoly of transportation. The owners of the diamond fields of Africa have acquired, by concerted action, a practical monopoly of the supply of diamonds.

Essentially similar is the monopoly which rests on the possession of some unique object or class of objects which has been produced in the past and which is incapable of reproduction; for example, the ownership of a statue dug up from the ruins of an ancient city or of all the copies of a rare first edition.

There is also the personal monopoly, which occurs when anyone is in position to render a service which no other person can provide. When the great pianist, Paderewski, plays in any city, there are always people who go to hear him, not because they can detect any superiority in his performance over any one of a hundred other pianists, but simply because they want to hear — or "to see" — Paderewski. When a distinguished statesman or poet goes on a lecture tour, he counts on a similar response. He may be a very poor speaker, he may have no very interesting message to impart, but there will always be persons who want to see him and hear him

speak. The shrewd politician is often, on account of his official position, able to render valuable services not obtainable from anyone else, and much "honest graft" is the cashing in on personal monopoly. Even the famous lawyer or surgeon has a monopoly position in so far as there are those who want his services to the exclusion of all others.

Capitalistic monopoly. Among the various forms of natural monopoly, the one about which the most fundamental economic problems center is the capitalistic monopoly. This type of monopoly rests on no exclusive control of the raw materials of industry, nor yet on any special privilege conferred upon it, although these also may be present and may facilitate its growth. It arises rather out of the highly capitalistic method of production which characterises modern industry, and a concern may both acquire and maintain its position of monopoly not by virtue of superior productive ability but because of more ample financial resources.

The foundation of capitalistic monopoly. The capitalistic monopoly is in large measure the by-product of large scale production. When production is on a small scale the number of actual or potential competitors is so great as to impede the formation of monopoly. Even if a monopoly be formed the capital outlay necessary to enable anyone to challenge the position of the monopoly is so small that the life of the monopoly is likely to be short. In agriculture monopoly has failed to appear to any great extent, not because of the disinclination of farmers to gain at the expense of their fellows, but because agriculture is carried on best on a small scale. There are consequently so many farmers that a union among them which would be sufficiently strong to control prices is almost inconceivable. In many lines of manufacturing however the superiority of the large producing unit is such that the small producers have been squeezed out and have left the field to a smaller number of large scale units, thus simplifying the problem of effecting a combination and ultimately securing monopoly.

In a still more significant way large scale production has tended in the direction of monopoly. Large scale production generally means a relatively heavy permanent investment. This not only

tends to discourage potential competitors but it has a determining influence on the type of competition among those already in the field. Cutthroat competition, as previously analyzed [1] and illustrated in the case of railroads is apt to develop in any industry where the fixed costs are heavy in relation to the operating costs, and the effects of competition of this type are so ruinous that even a threat of a price war may cause a scramble of competitors to seek shelter in combination.

Not only may competitors avoid the consequences of cutthroat competition by combining, but in some cases they may be able to effect economies in manufacturing and selling as the result of combination. If savings of considerable magnitude are the result of combination, a powerful economic motive exists for the formation of a combination, which may be a step toward monopoly.

The problem of capitalistic monopoly. It is principally to this type of monopoly that our attention will be directed during the remainder of the present chapter. The reader will doubtless recognize that complete monopoly — the exclusive control of supply — is seldom if ever attained by the mere capitalistic monopoly without the aid of some other contributing cause. It is virtually certain that some other sellers will always be present in the market. However the serious public problems of monopoly do not wait upon the complete consummation of capitalistic monopoly but appear as soon as one concern has progressed so far in that direction as to be able to shape its selling policy without regard to its small rivals and so to wield a predominant influence in the control of price. Indeed the problem may arise even earlier than this; that is, as soon as there is evident a clear tendency toward monopoly of the capitalistic sort.

The term "trust," although it has another and technical meaning which we shall consider hereafter, has in common language come to be identified with the large industrial enterprise of monopolistic power or tendency, and we shall find it convenient to adopt the term with that significance. The reader must be on his guard

[1] See Chapters VII and XXIX.

however against the looseness of popular usage. He will not confuse mere size with monopoly power. Some of the very largest industrial concerns are in active competition with rivals, perhaps equally large. Nor must combination be confused with monopoly. Though combination may frequently be a step toward monopoly, and though some combinations have monopoly power, there are many cases of large combinations in keen competition with each other, as for example the Carnegie Steel Company and the Federal Steel Company before the formation of the United States Steel Corporation.

The trust movement. In the United States the movement toward combination for the sake of monopoly power — " the trust movement " — dates from the late seventies; in the eighties it gathered momentum, and in the nineties it attained its greatest proportions. Some indication of the extent of the movement can be gathered from the statistics of combinations formed. In the period from 1887 to 1897 eighty-six combinations were formed, in 1898–1900 one hundred and forty-nine, and one hundred and twenty-seven in the years 1901 to 1903. Relatively few of the combinations formed have been " trusts," but from 1898 to 1900 a larger proportion of the combinations succeeded in acquiring monopoly powers in their respective fields than in any other period before or since. After 1903 the movement slowed down, partly because of the difficulty experienced in marketing the huge volume of securities issued at the time of the formation of each new combination, partly on account of adverse court decisions in anti-trust cases, and partly because of the fact that trusts had already been formed in most of the lines where monopoly was at all possible and indeed in some lines where monopoly power could not be retained.

Monopoly devices: the pool. Among the many devices used at one time or another to secure unity of action among competing producers, the pool is the oldest and probably the most frequently used. The term designates any sort of agreement between independent producers whereby their freedom of action in respect of certain matters is surrendered to a representative committee or to the group acting as a whole. Agreements to maintain prices, to

divide profits, to restrict output, to divide the selling field among themselves, to provide for the common use of all patents, and so on, are examples of pooling agreements common to many lines of industry.

One of the most interesting and most successful pools was the Wire Nail Pool organized by the Association of Wire and Cut Nail Manufacturers. The agreement provided for division of the output among the different manufacturers, fixing the total quantity to be sold each month, and setting the price at which the nails were to be sold. In January, 1895, the price of 8d wire nails was $1.45 per hundred-weight. Shortly thereafter the pool was formed, and the price was raised successively until in May, 1896, it reached the high point of $3.15, where it stayed for six months, when competition finally caused the disruption of the pool. The success of the pool is peculiarly interesting because of the fact that relatively little capital investment was necessary as a preliminary to manufacture. One of the men connected with the association throughout its existence attributed the success to its boldness in raising prices. "Suppose we had put up the price $.15 a keg. A great many men who had their eyes on the nail business would have said: 'These men have put up the price $.15, and they will probably make it $.25. There is going to be money in the nail business. I will go in.' But when we put up the price $1.50, they said: 'These fellows are lunatics. They can't hold together. There will be a smash before they get fairly started. I will keep out of that business.'"[1]

Pools have the advantage of retaining the identities and, so far as the methods of production are concerned, the independence of the member companies; furthermore they can be kept secret, and for that reason they are resorted to when other methods seem inexpedient. But, while some of them have been quite successful, they have generally not been continuous, stable, or long-lived. The independence of each member is bound to bring about a conflict of interest. The restriction of output, which is necessary to secure

[1] W. Z. Ripley, *Trusts, Pools and Corporations*, p. 64. Quoted from *Iron Age*, Dec. 3, 1896, p. 1108.

higher prices, means of course rising costs of manufacture, and some black sheep among the members may succumb to the temptation to keep costs down by manufacturing more than the amount allotted to him and disposing of it without the knowledge of his partners. Furthermore the pooling agreement has never had any standing in the eyes of the law. Even in the absence of legislation making it positively illegal, any agreement which unduly restrains trade is, according to the common law, void and unenforceable. While the common law exacts no penalty for making such an agreement, the authority of the courts cannot be invoked to compel the parties to the pool to live up to their agreement or to pay damages to the other partners for breaking the agreement.

The trust. The instability of the pooling agreement led to a search for some device which would be stable and at the same time legal. The *trust,* created first in 1879 by the Standard Oil interests, seemed to meet these requirements. The trust, as technically defined, [1] consists of the creation of a board of trustees, to which are transferred the shares of stock in the several companies, together with the assignment of voting power. In return the former holders of the stock receive trust certificates, which give the holders a proportionate share of the joint net earnings of all the stocks held in trust. Since the trustees hold a controlling number of shares of stock in each company, they are able to elect the several boards of directors and thus to attain uniformity of policy and action.

Forty separate oil companies were thus brought together by the Standard Oil interests. The plan commended itself to others and was immediately adopted in the cottonseed and linseed oil business, by the whisky distillers, by the sugar refiners, and others. But the formation of these trusts precipitated an anti-trust movement in the years around 1890. Various states passed anti-trust laws and attacked the trusts in the courts, and the passage of the Sherman Anti-Trust Act in 1890 indicated the awakening of the interest of the national government in questions of monopoly. In

[1] It will be noted that the term "trust" is used in two senses. From a technical viewpoint a trust is the form of combination here described, but the use of the term to describe any monopolistic combination whether in the form of trust or otherwise is too deeply imbedded in our language and thought to be rejected.

New York the North River Refining Company, controlled by the Sugar Trust, was attacked under the common law on the ground that it had exceeded its corporate powers in entrusting the direction of its affairs to an irresponsible board and that the trust was essentially a partnership, into which a New York corporation could not legally enter. As a result of this action the charter was revoked and the corporate existence of the company terminated. In Ohio the Attorney General brought action against the Standard Oil Company of Ohio, because of its connection with the trust, and succeeded in forcing the severance of the union. The effect of these decisions was to make the trust formed to restrain trade illegal under the common law.

The security holding company. The unfriendly reception accorded the trust led to the utilization of two other devices, the security holding company and the consolidation or merger of competing companies. The security holding company is a very simple form of organization. A corporation is organized to buy up a sufficient number of the shares of the companies which are to be combined to give it control in each case. Practically no capital is required to start the holding company, since the shares of the constituent companies may be obtained in exchange for the stocks or bonds of the holding company. As in the trust, unity of policy is obtained through the election of the proper men to the boards of the constituent companies. The development of the security holding company had been delayed by failure on the part of state legislatures to give to corporations the power to hold the stocks of other corporations. In 1889 and 1893 however New Jersey amended its corporation laws so as to permit corporations to hold stock in other companies and to exercise the right of owners, and the fear lest New Jersey become the home of all large corporations led some of the other states, notably Maine, New York, and Delaware, to liberalize their laws in the same fashion.

Consolidation. At this time, as a matter of fact, consolidation was more frequently resorted to. In this case the corporation acquires the property of the companies which are to be combined, either by consolidation and merger or by the purchase

of the property for cash or for the stock of the holding company. The American Tobacco Company, formed in 1890, exchanged its stock for the plants, business, brands, and good-will of five companies, and the American Sugar Refining Company, organized in 1891, was another example.

Advantages and disadvantages of holding companies. The security holding company has certain advantages over the consolidation as a form of organization. It is easy to form, whereas there may be many technical difficulties to prevent a company from merging itself in another or even disposing of its plant and equipment. The security holding company also preserves the identities of the individual companies and consequently the good-will which may attach to the corporate name, and it can adapt itself to the requirements of the individual states, by incorporating a subsidiary in each state if necessary. On the other hand it is a complex form of organization; in contrast to the concentration of responsibility and authority in the consolidation, there must be a complete administrative system for each member company. Neither form of organization has however proved to be legal when formed for the purpose of monopolizing trade.

Anti-trust legislation in the United States. American public opinion has always looked askance upon monopoly and indeed upon all large business enterprises and combinations which bear any odor of a tendency toward monopoly. In the course of the rapid movement toward combination whose outline we have sketched, there arose fear lest the entire industrial fabric become tainted with the virus of monopoly and the control of industry become concentrated in the hands of a relatively small number of men. One result was the enactment of legislation, both by the states and by the federal government, designed to break up trusts and to prevent the existence of conditions favorable to the development of monopoly. Brief reference has been made to state legislation and state prosecutions. Our further attention will be given to the activities of the federal government, which, on account of its jurisdiction over interstate commerce, has been most directly concerned in the political attacks upon the trusts.

The Sherman Anti-Trust Act of 1890. With the exception of the Interstate Commerce Act, no legislation designed to control trusts was enacted until 1890, when the act commonly called the Sherman Anti-Trust Act was passed.

The Sherman Act is brief and exceedingly general in its phraseology. In the first section "every contract, combination in the form of a trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States or with foreign nations" is declared illegal, and according to the second section "every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a misdemeanor. . . ." Under the act the government may bring criminal prosecutions against persons guilty of violating its provisions, or it may apply to the circuit courts, by means of suits in equity, to prevent and restrain violations of the act. Upon conviction in the courts of a violation of the criminal sections of the act, a penalty, consisting of fine or imprisonment or both, may be imposed. Any person injured in his business or property by any act declared unlawful or forbidden by the Sherman Act is authorized to start suit in a circuit court of the United States, and if his claim of damage be upheld in the court he can recover three-fold the damages sustained and the costs of the suit including a reasonable attorney's fee.

The operation of the Sherman Act. The criminal features of the Sherman Act have not proved especially useful in preventing the growth of trusts, so difficult has it been to secure a conviction by jury on the rather vague charge of restraint of trade. On the other hand in the field of forcing the dissolution of trusts, the government has met with greater, though by no means unvarying, success.

The first case of importance was United States *v.* E. C. Knight Company, decided by the Supreme Court in 1895. In this case the government asked that the purchase of the E. C. Knight Company and three other companies by the American Sugar Refining Company be set aside on the ground that the contracts of purchase constituted combinations in restraint of trade. The court however

refused to accede to this, basing its decision on the failure of the government to offer proof of any direct restraint on interstate trade resulting from the purchases. Actually the case was lost by the inadequate and faulty presentation of the case by the Attorney General; nevertheless the decision served to foster the belief that the Sherman Act was ineffective and to encourage the formation of trusts.

To the indifference or neglect of those responsible for the administration of the Sherman Act can in large measure be attributed the fact that from 1890 to 1901 there were but eleven bills in equity, five indictments, and two contempt proceedings brought under the act, few of which were of any consequence or attended with success. The most notable was United States *v.* Addyston Pipe and Steel Company, instituted to force the dissolution of the cast iron pipe combination and decided in 1899. In this instance the courts upheld the government, and the falling off in the formation of trusts after 1899, which we have already noted, may in part be attributed to the proof thus given that the Sherman Act was still alive.

President Roosevelt began the energetic campaign against the trusts. In the years of his administration (1901–1909) eighteen bills in equity, twenty-five indictments, and one forfeiture proceeding were instituted. Moreover new legislation was passed to facilitate the handling of anti-trust cases. We have already noted the Expediting Act (see Chapter XXX), which gave precedence in the federal courts to cases arising under the Anti-Trust Act and the Interstate Commerce Act. Another act secured the creation of a Bureau of Corporations, designed to give the government and the public authoritative information regarding the activities of the trusts.

One of the first of the significant cases to come up at this time was the Northern Securities case.[1] The Northern Securities Company had been organized as a New Jersey corporation to acquire control, by stock ownership, of two parallel and presumably competing rail-

[1] The government had already (1897) established the point that railroads were amenable to the Sherman Act.

roads, the Northern Pacific and the Great Northern Railway Companies. The government instituted a suit in equity to have the company dissolved, and in 1904 the Supreme Court rendered its decision ordering that this be done. Of still greater importance was the dissolution of the Standard Oil Company which was ordered in a decision rendered in 1911. This case marks a new interpretation of the terms of the act. According to the language of the earlier decisions all combinations or contracts in restraint of trade were held to be unlawful, but in this decision the Supreme Court enunciated the so-called "rule of reason." As under the common law, so under the Sherman Act, a restraint of trade which was not undue or unreasonable and which was incidental to a lawful purpose was henceforth not unlawful.

We may note one other case, that involving the Tobacco Trust. The Tobacco Trust was in part formed by consolidation, and it remained to be seen whether outright consolidation of the competing businesses would stand the legal test. The court however went behind the form to the motives and results of the activities of the Trust and ordered its dissolution.

The anti-trust legislation of 1914. While considerable success had attended the efforts to dissolve trusts under President Roosevelt — and President Taft continued the program with equal energy — the feeling was widespread that more rigorous legislation was desirable and that an administrative body should be created with greater powers of regulation and investigation than those possessed by the Bureau of Corporations. This feeling was clearly reflected in the campaign pledges of the three parties actively in the field in the election of 1912. Mr. Wilson came into office pledged to "the proposition that private monopoly is indefensible and intolerable," and early in 1914 he called upon Congress to enact legislation to supplement and amend the Sherman Act. As a result of his recommendations two new laws, the Trade Commission Act and the Clayton Act, were passed by Congress in 1914. The interest of the public in curbing monopoly is evidenced by the fact that the first act was passed on a non-partisan basis and that the Clayton Act had the support of many of the Republican members of Congress.

The Trade Commission Act. A Federal Trade Commission, consisting of five members appointed by the President with the advice and consent of the Senate, was created by this act. Each member was to receive a salary of $10,000 a year and, save for the first appointees, to serve a term of seven years. Not more than three members might be from the same political party. With the organization of the Federal Trade Commission the Bureau of Corporations ceased to exist, and all its pending investigations, records, and property were taken over by the Commission.

Investigation of corporations. The Federal Trade Commission was given comprehensive powers of investigation and report, and to facilitate its work it was granted the authority to call upon business men to produce records of all sorts and to give testimony when required. It may investigate the practices of any corporation guilty or alleged to be guilty of violating the anti-trust laws; furthermore it is permitted, and in some circumstances required, to investigate the way in which a corporation is carrying out the decree of a court in an anti-trust case. But its powers are not confined to the investigation of infractions of the anti-trust laws, for it may call upon any corporation for information as to its business organization, methods, and relations with other business units. This provision was not inserted in order to harass business men, but to enable the Commission to accumulate a fund of general knowledge with regard to the conduct of business which would be of value to Congress or to the public.

The field of business research was not confined by the law to the United States; a specific part of the duty of the Commission is to keep Congress informed of trade practices in foreign countries which may influence American foreign commerce, and one of the earliest reports issued by the Commission was on the subject of combinations in foreign trade with special reference to their effect on American export trade.[1] This report was incidentally a factor of considerable importance in securing the passage of the Webb-Pomerene Act which is described on page 65.

[1] *Report on Co-operation in American Export Trade*, 2 vol., 1916.

Advisory functions. Certain advisory functions were also provided for by the Trade Commission Act. It was felt that the Federal Trade Commission, as a result of its investigations and research, would be peculiarly fitted to assist the Department of Justice and the courts in an advisory capacity. Many of the questions which arise in connection with anti-trust cases are primarily economic, and it is reasonable to suppose that a body of economic experts may supplement the work of legal experts to great advantage. With this view in mind it was in the first place made the duty of the Commission, upon the application of the Attorney General, to recommend a method of readjustment which would bring within the law a corporation alleged to be guilty of violating the anti-trust laws. An amicable settlement was believed in some cases to offer advantages over a long-drawn out and costly law-suit. In the second place the Commission was required, at the request of the court, to prepare an appropriate decree in any suit in equity brought by the Attorney General under the anti-trust laws. The court might accept the decree, modify it, or reject it entirely, but at least it would have the benefit of expert advice. Finally, the Commission was given the right to recommend legislation to Congress on the subjects which lie in its field.

Unfair competition. It is also the duty of the Federal Trade Commission to prevent unfair methods of competition. In cases adjudicated under the Sherman Act, the courts had often been influenced in their decisions by acts which, while perhaps innocent in themselves, could not be considered lawful when practiced by a great combination in circumstances which seemed to indicate an intent to destroy competition rather than merely to gain business. Among the many methods which have been held to be unfair, we may mention: (1) price discrimination, which means that prices are cut, below the cost of production if necessary, at the competitive points while being maintained at a high level elsewhere; the Tobacco Trust and the Standard Oil Company were notorious in this regard; (2) the exclusive dealers' or factors' agreement, which compels a wholesaler to handle the goods of one concern exclusively or penalizes him, by charging him higher prices, if he does carry

competitive goods; (3) the use of "fighting" brands, especially designed to be sold at a price ruinous to competitors; (4) brand imitation; the American Tobacco Company thus put out at one time a "Central Union" to kill the sales of the "Union Leader" of a competitor; (5) espionage on the customers, methods, shipments, etc., of a competitor; and (6) organization of companies ostensibly independent to win the confidence of competitors.

Unfair methods of competition were not considered legal under the Sherman Act, and the Trade Commission Act adds nothing to the law in this respect, but it does create a method whereby relief may be obtained without the cumbersome recourse to a damage suit, which alone was possible previously. Now a concern which believes that it is being injured by an unfair method of competition may complain to the Federal Trade Commission; if the Commission is convinced of the justice of the complaint, it serves upon the offender a complaint, giving the charges. A hearing is then held, at which time the one on whom the complaint is served has an opportunity to show why a restraining order should not be issued. If the Commission still feels convinced of the justice of the complaint, it issues a restraining order to the offender, and if the order is not obeyed, the Commission may apply to the Circuit Court of Appeals for enforcement of the order. The court then has complete control over the case and may confirm, modify, or set aside the order of the Commission. The offender also has the right to petition the court to set aside the order, and he has the further right of appeal to the Supreme Court.

The Clayton Act. The Clayton Act prohibits, with certain limitations, local price discrimination, tying contracts, holding companies, and interlocking directorates, the intention being, by making these practices specifically unlawful, to stifle the creation of trusts, conspiracies, and monopolies in their incipiency and before their consummation.

It is made unlawful for a person engaged in commerce to discriminate in price between different purchasers of a commodity, where the effect of such discrimination may be substantially to lessen competition or may tend to create a monopoly in any line of

commerce; except that discriminations may be made for differences in the grade, quality, or quantity sold, or in the cost of selling or transportation, or when made in good faith to meet competition, and that a seller may select his customers in bona fide transactions and not in restraint of trade.

Tying contracts, which bind the purchaser or lessee of certain products to buy or use certain other products exclusively, are declared illegal where the effect of such lease, sale, or contract for sale may be substantially to lessen competition or may tend to create a monopoly in any line of commerce.

No corporation engaged in commerce may acquire, directly or indirectly, the whole or any part of the stock of another corporation also engaged in commerce, where the effect of such acquisition may be substantially to lessen competition between the two corporations, or to restrain such commerce in any section of the country, or to tend to create a monopoly in any line of commerce. Holdings of securities already acquired are not made illegal, and a corporation may continue to acquire securities for investment or other purposes, provided it does not offend in the above particulars.

No person may at the same time be a director in two or more corporations engaged in commerce (other than banks and common carriers), any one of which has capital, surplus, and undivided profits exceeding a million dollars, if such corporations are or have theretofore been competitors so that the elimination of competition by agreement between them would constitute a violation of any of the provisions of any of the anti-trust laws.

Remedies provided by the Clayton Act. Certain remedies were provided by the Clayton Act covering the violation of its provisions or of those of other anti-trust acts. We may conveniently summarize them in four groups:

(1) The Interstate Commerce Commission is charged with enforcing the provisions when they apply to common carriers, the Federal Reserve Board when they apply to banks and trust companies, and the Federal Trade Commission when they apply to other corporations; and the manner of procedure is specifically outlined.

(2) Any person injured in his business or property by reason of anything forbidden in the anti-trust laws may institute suit and recover threefold damages, together with costs of the suit. If in a suit brought by the United States under the anti-trust laws it is decided that the defendant has violated said laws, this fact must be accepted as *prima facie* evidence against the defendant in any suit brought by another party against him under the anti-trust law. This provision makes it easier and cheaper for the injured person of small means to sue a trust.

(3) The district courts of the United States are invested with jurisdiction to prevent and restrain violations of the Act, and the district attorneys are charged in their respective districts, under the direction of the Attorney General, to institute proceedings to enforce the Act.

(4) Any person or concern is entitled to injunctive relief, in any proper court of the United States, against threatened loss or damage by violation of the anti-trust laws, a preliminary injunction being issuable when it appears that irreparable damage is immediate. This makes it possible for an independent concern to defend itself against a trust, when attacked by some unfair device, and to maintain its existence, whereas it was formerly possible only to sue for damages after injury and perhaps bankruptcy had resulted.

The Webb-Pomerene Act. In 1918 the trust policy of the United States was modified by the passage of the Webb-Pomerene Act, which exempted from the provisions of the anti-trust acts combinations engaged solely in the export trade provided such combinations do not restrain trade within the United States nor employ unfair methods of competition with American competitors. This liberalization of the anti-trust laws was conceived to be necessary to enable our exporters to compete with those of countries where combinations are either tolerated or encouraged. The result has been to allow small manufacturers, who otherwise would find it difficult to enter the foreign markets at all, to enter as a combination. The larger manufacturers have shown little inclination to form export combinations, and it is doubtful whether the

act will ever succeed in stimulating our export trade to the degree anticipated by its sponsors.

Conclusion. The government regulation of trusts has in many respects been but partially successful, even in the years when the dissolution of trusts was being most energetically and successfully pushed. To dissolve a trust is not the same thing as to restore competitive conditions. If the dissolution decree leaves a small number of stockholders in control of the different companies, community of interest may succeed where formal organization has failed. Then too the courts have failed to give full recognition to the fact that the dissolution of a trust is an economic as well as a legal question. It was expected that the Federal Trade Commission would be of great assistance to the courts in framing the dissolution decrees, but thus far the courts have not made extensive use of the services of the Commission.

Nevertheless substantial progress has been made, and it cannot be questioned that overt inclination on the part of combinations to reach out for greater power or to crush competition by questionable means is less than formerly. It has been demonstrated that the anti-trust laws have teeth — that they can be effective if need arise.

Various proposals have been made to improve the existing laws and to substitute an entirely different method of regulation. In order to secure uniformity in the powers given to corporations engaged in interstate trade, some students of the problem advocate federal incorporation or, if the constitutionality of this be questioned, a federal license for all such corporations. This would remove the difficulty caused by the diverse regulations of the different states. It has also been proposed that one who owns stock in several competing corporations be permitted to choose the company in which he shall exercise the right of voting. This presumably would make it impossible or difficult for a small group of men to suppress competition by acting together.

A more radical proposal would substitute the regulation of prices for the regulation of conduct. Let monopolies be formed, if the owners desire, but safeguard the interests of the public by fixing the

prices which can be charged. The difficulties which would attend the practical operation of such a scheme remove it, for the time being at least, from the plane of practical proposals, and the same holds true of the proposition for the nationalization of all industry. Our present method of regulation has met with a sufficient degree of success to make a radical departure seem unwarranted. Nevertheless it should be recognized that the failure of the present policy, in whole or in part, may bring recourse to more drastic methods.

Is monopoly inevitable? The many forces which appear to be tending toward monopoly and the apparently limited power of the public to prevent monopoly by restrictive legislation raise the interesting and far-reaching question whether there is after all a fundamental tendency toward monopoly which it is idle to resist — whether in short monopoly is inevitable.

If we turn back to envisage the whole field of monopoly of every sort, it becomes at once apparent that monopolies of the first group, the public monopolies, are not inevitable. They are created by law; the government may if it choose cease creating them, and it may eventually (with possibly a few exceptions) terminate the existence of those already created.

But what of the natural monopolies? Evidently those monopolies which rest on the existence of a sole natural source of some product or upon the existence of unique non-reproducible things are inevitable. Their present owners may be dispossessed, but the monopoly, even though ownership be vested in the government, will continue. Personal monopolies also are inevitable.

But the real problem of monopoly is not much concerned with legal monopolies or with these first three types of natural monopoly. The problem has to do with the fourth type, the capitalistic monopoly. If it be true that natural laws, operating in a social system based as ours is upon personal liberty, private property, and individual control of enterprise, tend inevitably to bring about a condition of monopoly in the production of a substantial part of the basic products which mankind requires, then we are faced with a human problem of the first magnitude, a problem which may well put upon the defensive the very fundamentals of the existing social order.

Forces favoring monopoly. At first glance the forces favoring monopoly appear very strong. There is first of all the well known economy of large scale production in many lines of industry. There are the economies which in certain cases appear to accompany the combination of separate business units under a single management. There is the occasional appearance of cutthroat competition, an unstable situation leading almost inevitably to monopoly. There is the constant effort of the captains of industry, when competition becomes especially keen and irksome, to find escape through some form of combination, making use of every possible device — pool, trust, holding company, and so on — which human ingenuity with billions at stake can devise. The corporate form of organization is a helpful instrument. The protective tariff, while probably not the " mother of trusts " as it has sometimes been called, has unquestionably reduced the competition which American producers have had to face and in that respect has coöperated with the other forces tending toward monopoly.

Acquaintance with economic history will convince one that the enactments of legislative bodies are futile if they are in opposition to powerful economic forces. If cutthroat competition is an inevitable feature of large scale production, it would be the part of wisdom for society to seek stabilization in monopoly and to direct its efforts to regulation calculated to enforce good service and reasonable prices. If a combination of producing units under one management is more efficient than when each of the units is under separate management, if a combination of combinations in exclusive control of the field is the most efficient of all, it is folly to try to run counter to the natural tendency of business men to seek the most profitable form of organization. Society cannot permanently reject the most economical type of productive organization. This is virtually the conclusion which we have reached as regards the railroads, and our railroad policy and legislation are based largely on the assumption of the inevitability of monopoly.

Limits to monopoly. But the conditions which prevail in the realm of transportation, communication, and the municipal public

utilities, which confessedly point the way to ultimate monopoly, are not universal throughout the whole field of production.

At the forefront stands agriculture, most basic of all basic industries, where economy is not to be found in large scale production or combination, and where fundamental economic laws seem destined to preserve indefinitely the régime of small productive units and free competition.

In the fields of manufacturing and marketing it is doubtful if the monopolistic combination is in many cases the most efficient form of organization. We have already learned that the efficiency of large scale production and combination has very real limits. There are few important lines of industry in which this limit would not be reached long before the would-be monopolist had become great enough to absorb the whole. Even assuming the monopoly to become established, lethargy in management and lack of progress in the technique of production and the methods of organization will tend to appear, so that the monopoly which at the moment of its formation is representative of the most efficient methods of production carries with it no assurance of continued efficiency.

There is very little reason, from the standpoint either of theory or of history, for believing that the economies of combination are sufficient in volume to warrant the growth of monopolies in many fields of production. Nor is there evidence to show that competition between large units must necessarily be of the cutthroat variety. This is essentially the result of an unbalanced condition in production, where the total productive capacity in a given industry is greater than the market requires. When the intense competition has eliminated enough of the producers to enable those remaining to work to capacity, the incentive to lower prices in order to attract business is removed, since increased production will no longer bring lowered costs of production. The industry may then settle down to a normal condition of competition between a few strong rivals, no one of which has reason to hope for monopoly power.

There remains the problem which is created by the monopolistic ambitions of certain producers or group of producers. While it

may not be possible for a manufacturer to reduce his producing costs by drawing trade from his rivals, it may be worth his while, if he be strong enough financially, to sell for some time at a loss with the expectation of more than recouping his losses when his competitors have been driven into bankruptcy and he is left in control of the field. Under the guise of competition he may be dealing destruction to rivals who are weaker only in the sense that their resources are smaller, making them vulnerable to the insidious attack of the would-be monopolist, who brings to his aid, not merely the weapon of price cutting, but the many devices known as " unfair methods of competition," by means of which the concern with more ample financial resources can undermine the position of a financially weaker rival.

So deeply ingrained is this desire to secure monopoly profits that we may safely say that it can never be eradicated, but there is reason to believe that here the power of organized society, acting through legislation and administration, is strong enough to cope with the situation. As we have seen, progress has been made toward check-mating the attacks of those enterprisers who are not willing to " play the game " according to the rules. With such influences eliminated, there is good reason for the conclusion that, throughout the greater part of the industrial field, fundamental economic forces are working, not to bring about monopoly, but rather to preserve a state of active competition, none the less intense because it may often be carried on by a relatively small number of powerful rivals.

CHAPTER XXXII

RISK, INSURANCE, AND SPECULATION

Futurity, human knowledge, and risk. The most important attribute of the present moment is that it is the dividing point between past and future, a cleavage whose profound significance must never be forgotten or misunderstood by the student of economics. The past is gone; nothing can be done to change it. The future is before us, and it is for the sake of future effects — pleasures to be secured and pains to be avoided — that all economic activity is put forth.

Now while personal experience and recorded history tell us with more or less certainty of past events, the future is notoriously uncertain. Man cannot say that he is absolutely certain of any future event, though the degree of uncertainty varies all the way from the almost complete certainty that the sun will rise tomorrow to the complete uncertainty as to what the direction of the wind will be a year from today. Thus the very nature of the universe in which man lives compels him to bear *risk; i.e., the chance of loss or injury.*

Since the uncertainty of future events permeates every phase of human existence, risk is universal. It takes a thousand forms — the uncertainties of death, disease, unemployment, fire, change of price, defective title, rise of costs, change of fashion, strikes, spoliation, piracy, tariff changes, tornado, earthquake, political instability, strength of materials, declaration of war, and so forth without end.

It is evident that the basis of risk is the lack of human knowledge of future events. To an omniscient person, who could foresee with perfect accuracy the future, there would be no risk. To actual human beings risk is always present, because no one is omniscient,

and it is small or great according to the degree of our knowledge. The risk of loss from frost either in August or in January is almost *nil*, because of our well-nigh certain knowledge that that event will or will not occur as the case may be. Owing to our relative ignorance of the event, the risk of loss by hail in August is much greater.

The human attitude toward risk. The instinctive reaction to the unknown is fear, and it is no cause for wonder that primitive man, beset with manifold dangers of whose origin and nature he was ignorant, peopled the universe with malevolent spirits whose activity might be allayed and who might be propitiated by ceremonials and sacrifices. Modern man takes a calmer and more rational view of risk. He has ceased to regard future uncertainties solely from the standpoint of superstition, though his emancipation is not yet complete — witness the common though shamefaced dread of thirteen at table or a wedding on Friday.

In his attitude toward risk, man displays a mixture of emotions. A certain amount of risk now and then is not unwelcome to most of us. Not content with all the innumerable risks which nature provides, we go out of our way to create other hazards. This is *gambling*, the intentional creation, for its own sake, of artificial and unnecessary risk. The man who makes a wager that his football team will win a particular game has created for himself an entirely new risk. Of course certain risks — of chagrin, loss of prestige, etc. — attached in any case to the outcome of the game, but the gambler has added the risk of losing money by the defeat of his team. Gambling such as this — between honest amateurs, with the odds fairly adjusted to the chances — is indulged in because of the pleasurable excitement of having something at stake; because men like " to take chances." Even when the odds are somewhat against the gambler, he may consider the amusement and excitement obtained, as at the gambling table or the races or the lottery drawing, a sufficient return for the losses which in the long run he is likely to suffer. The professional gambler's attitude is different; to him it is a cold-blooded business proposition, indicating no pleasure of taking chances, except on the part of his clients.

But while a certain modicum of the gambling spirit is well-nigh universal, men do not relish too much risk, and in general they regard the risks which nature imposes upon them as an evil, to be avoided or reduced wherever possible. In this contest against risk man has a variety of weapons to his hand.

Risk reduced by knowledge. The essence of risk being ignorance, it follows that just so far as knowledge supplants ignorance risk is reduced. There are certain classes of events which are wholly beyond the control of man; yet if by the advance of scientific knowledge he becomes able to predict their occurrence with some approach to certainty, he is enabled to take appropriate action to prevent or minimize his chance of loss or injury. Man can do little or nothing to prevent the occurrence of storms, floods, earthquakes, and other natural hazards, and our present knowledge of seismic and meteorological forces is so meagre that earthquakes and weather defy forecasting much in advance. Yet the United States Weather Bureau, in spite of the fact that its forecasts are not altogether dependable even for twenty-four hours in advance, gives information which makes possible avoidance of much loss by provisions taken against forecasted storms.

As scientific research adds to human knowledge, prediction becomes steadily more accurate and risk is correspondingly reduced. Of recent years remarkable strides in this direction have been made in the business world. Entrepreneurs are today inclining less and less to take an order, hire an employee, locate a store, write an advertisement, despatch a salesman to a territory, or offer a security for sale, without a preliminary investigation of all available data, so as to reduce the risk of a snap judgment based on imperfect information. The business man, through his commercial research department and statistical and advisory bureaus, is fast becoming expert at analyzing and synthesizing economic conditions and directing his course in less risky ways.

Preventive measures. A second method of dealing with risk is to prevent the occurrence of the harmful event. The construction of fireproof buildings, the installation of fire-fighting apparatus, the establishment of safety deposit vaults, the provision of free

medical advice and care of insured people, the institution of devices to stabilize the price level — all these are instances of the reduction of risk by preventive measures. New methods of prevention are being constantly devised, and the costs of many of the old methods are being reduced.

This method of dealing with risk is intimately associated with the former, for preventive methods cannot be devised until scientific investigation has unearthed the cause of the event. Medical science offers many illustrations. Not many generations ago the risk of death by small-pox was appallingly great, now the discovery of the vaccine against small-pox has greatly reduced the ravages of the disease. Curiously enough the success of the preventive measures has induced an underestimate of the danger of postponing or neglecting vaccination, so that the risk of an epidemic of small-pox is in some localities greater now than in the past.

(3) **Combining risks: the law of large numbers.** A third method of eliminating risk is to combine risks by operating on a large scale. This is based on the principle of scattering risks, of not "putting one's eggs in one basket," or *the law of large numbers;* namely, that *a high degree of certainty may attach to a composite of data of which every component is highly uncertain.* Although one cannot know absolutely whether Jones will desert the army, experience will show about what number of men will desert an army of 100,000 men. Of a large group the percentage of men who will die from disease, fail in business, be killed on the railroads, go to college, etc., can be quite closely told in advance. An investor by scattering his funds among a hundred securities makes the risk of loss by default of debtor far less than it would be if his investment were all in one kind of security. A manufacturer who produces a variety of goods or who markets his product over a broad area with many customers faces less risk of loss of his market than one who clings to one product or to a local market with few buyers. By basing business forecasts and operations on averages of many factors, the risk factor is reduced and in some cases almost entirely eliminated. The variations of the numerous factors

counterbalance one another about the average. In combinations of competitors all those risks arising from ignorance of the probable action of a competitor are also eliminated; one of the chief arguments for large-scale industrial combination is the higher stability of business and the consequent lower risks under it.

Reserves. It has long been a commonplace expedient in banking to keep on hand a cash reserve against the contingency of extraordinary sudden demands by the customers of the bank. So long as funds synchronously flow to the bank in amounts equal to the withdrawals, no cash reserve is necessary, but it is the part of wisdom to avoid suspension of payments by carrying some idle cash against the contingency of withdrawals exceeding receipts. The cost of thus eliminating the risk of suspension is measured by the sacrificed interest on the reserve. The expediency of carrying reserves applies to all kinds of business, not to banking alone. To keep his plant going continuously the manufacturer maintains an inventory of raw materials and equipment in excess of the absolute amount which would be required if he could depend upon a regular flow of materials to his plant and continuous service from all his equipment. Likewise he carries a surplus of finished products against the possibility of extraordinary demands from old customers or the demands of new customers. The large corporations find their offerings of securities are much more readily sold if they have had a long record of steady dividends, and to that end they carry dividend reserve funds to make up the deficiencies of lean years and enable them to declare the regular dividends. In such ways is business stability enhanced and risks reduced by the maintenance of reserves.

It is important to note that by combination the absolute amount of such reserve funds may be reduced and the same reduction of risk be attained. For illustration, suppose it were statistically determined that a tank of 1,000 gallons of water is the minimum of safety to protect a warehouse of 1,000,000 bushels of wheat. If each of ten such warehouses kept its own tank of water, 10,000 gallons would be carried. But it is quite improbable that more than one of these warehouses would catch afire at the same time;

therefore, if water mains connected one common tank with all the warehouses, a reserve of 1,000 gallons would give as great protection. This is the principle of municipal fire protection. It is the principle of centralized banking, in which the reserves of the member banks are largely concentrated in a central bank which stands ready to throw the reserve funds to the needy bank undergoing a run. It is furthermore a basic consideration in industrial combination and large-scale operations, for the total quantity of reserve supplies may be much smaller when one plant feeds another plant with materials and one managing head coördinates their operations. In fact, as our economic society becomes more highly organized, each part carefully articulated with the others and avenues of communication and facilities of transportation well developed, the total capital held as reserves against all sorts of business contingency may be reduced with no increase of risk.

Passing risks on. In spite of all possible devices for eliminating and reducing risk, risk remains. Human knowledge can never attain the goal of complete prediction. The future will always be uncertain. And no human measures can wholly prevent untoward events with their consequent loss and injury. Risk is inevitable. Under these circumstances, expedients have developed whereby those who are unable or unwilling to bear the inevitable risks of life may pass them on to others more capable and more resourceful.

Specialists in risk bearing: the entrepreneurs. The simplest example of the passing on of risks is the relation between the laboring class and their employers. As we are fully aware, production in the modern world is involved from beginning to end in risk. The risks are inevitable, and someone must bear them. But they are not, as might at first thought seem natural, shared equally by the several parties to production. The laborers are not able to bear risks. They must have some assurance of a regular income, not dependent upon the success of any business venture; otherwise they would perish. This assurance is one of the characteristics of the wage contract; the employer relieves them of risk. Landowners and capitalists also, though not in the same degree

as the laborers unable to bear risks, nevertheless choose not to bear any great hazards, and they also pass the risk on to the entrepreneurs.

The entrepreneur hires laborers for a definite wage, he borrows capital at a specific interest rate, he buys or rents land at a fixed price or rental. If at the end of a period his gross earnings more than cover his costs, he enjoys a residue as his own wages of management and as profit. Otherwise, he fails to receive any share of the product and may even suffer loss. He thus becomes the chief risk bearer, enabling the other participants to avoid the greater part of the risks of production. For this service of risk-bearing, the entrepreneur's reward is profits, a form of income which fluctuates widely according to the outcome of business.

The nature of these several shares in the distribution of the social product and their relations to each other will be more fully investigated in a later part of this book (Part V). The point that we should grasp here is that the members of society automatically differentiate into groups on the basis of their willingness and ability to bear the risks of production and distribution, and that one class, variously called owner-managers, business men, or entrepreneurs, offers to take control, bear the chief risks, and appropriate what profits, if any, may result. The economy for society lies in the fact that, through this transfer to men who from the very fact of business leadership are presumably most competent to handle uncertain undertakings, risks are reduced. This competency consists in the possession of special ability, greater information and experience, sound judgment, self-confidence, and the willingness to undertake the risks of business and abide by the consequences. It is merely another instance of the division of labor. In the hands of this specialized class of risk bearers, productive enterprise is likely to be conducted with less risk of loss than if the hazards were more generally diffused.

The class of entrepreneurs is itself subdivided into classes for accomplishing a further specializing of the risk bearing function. By means of sub-contracts, risks are subdivided among those most competent to bear them in the several special fields. For instance

a builder, having secured the contract to build a house for a specified sum and wishing to protect himself against contingencies about which he may know relatively little in the structural steel, stone, lumber, cement, brass, plumbing, and other trades, immediately contracts with members of these trades to supply these materials at definite prices, thus shifting the burden of risks in these lines to the sub-contractors. By this process not only the risk of the transferrer, but also the total risk of society, is reduced, because the sub-contractor is presumably wiser in his specialized line than is the general contractor.

Insurance. The most important means of transferring risk is insurance. Insurance is based on the law of large numbers. Let us assume that in a certain community there are 1,000 house owners, the houses being worth on the average $10,000 each. Each owner bears the risk of having his house destroyed by fire at a loss (on the average) of $10,000. Let us suppose further that experience shows that in this community there is an average loss by fire of one house a year. Then if each owner would contribute ten dollars each year to a common fund, it would be possible out of this fund to reimburse anyone whose house was burned. In return for a small annual payment the householder is relieved of the risk of loss by fire. Risk is passed on by each one of the group to the group as a whole.

This is an example of what is called *mutual insurance*, in which the insurers and the insured are the same. If losses are sustained which are greater than the fund accumulated against them, each one of the group is liable for his proportionate share. Insurance may also be conducted by a corporation, which assumes risks as a matter of business in return for the payment of annual premiums, and whose shareholders expect the business to show a profit upon the capital which they have invested in it.

Insurance reduces risk. Insurance can be applied to any kind of hazard, provided a large enough number of independent risks can be combined and provided that the probability of the event insured against can be actuarily determined. When these conditions obtain, it is possible for the insurance company to predict

very accurately the number and amount of losses which it will have to bear in a year, and it can calculate the rate of premium which it will have to charge in order to meet these losses, pay expenses, and obtain a profit. Insurance thus applies the principle of combining risks and so reducing risk. In fact the total losses may be so regular and so predictable that the insurance company bears very little risk. While it may be highly uncertain what individuals aged nineteen will die in the United States in any year, on the basis of past experience it can be predicted with great certainty how many individuals of the age of nineteen will die during the year. The insurance company in writing life insurance policies for one hundred thousand people is not assuming the sum of the individual risks. In fact its risk from deaths is negligible, for it knows how many deaths to expect each year. The chief risk which it runs is that of failing to secure enough individual risks to make it possible to apply the law of large numbers. Note that it is not the primary function of insurance to prevent loss. There may be as many houses burned and as many deaths as if there were no insurance. Insurance does not eliminate loss; but it does reduce or eliminate risk.

When the two conditions laid down in the preceding paragraph are not fulfilled, insurance in name may be nothing more than gambling in fact. There are insurance companies or underwriters which will insure against practically any future event, as the death or illness of some particular person, rain on a particular day, etc. In such cases, risk is shifted to others, but without any real reduction in amount. Insurance against drought or hail or bad weather and other similar forms of insurance are hardly out of this speculative group as yet.

To the individual insurance appeals as a device whereby he can substitute small and certain payments (called premiums) for a highly uncertain and perhaps disastrously large loss. Without insurance many a manufacturer would be forced into bankruptcy if a fire were to sweep away his buildings, machinery, and materials, whereas there are but few who cannot afford the premiums in return for which an insurance company will assume the risk.

The manufacturer therefore pays the premiums, and as he expects to recover them in the price of his goods he pays them cheerfully. The individual likewise regards the payments for life insurance, automobile insurance, and the like as among the costs of living, infinitely to be preferred to the assumption of the risk of leaving his family unprovided for through sudden death or to the chance that he may face a ruinous damage suit as the result of an automobile collision.

Insurance may reduce loss. Insurance is therefore a stabilizing factor in private and business life, a convenience to individuals, and a reducer of risk. Moreover insurance may even exert an influence toward reducing losses. The insurance company has of course a direct interest in reducing the losses against which it insures. And the insurers, being by training and experience specialists in their particular lines, are competent to devise measures best adapted to the avoidance of loss. Thus a company which guarantees titles to real estate is in a position to have more accurate knowledge of the situation than the purchaser of a plot of land, and it may be able to prevent losses from bad or uncertain titles. Fire, marine, and casualty insurance companies impose elaborate rules and maintain inspection services in order to prevent fires and accidents. Even the life insurance companies have, in certain instances, inaugurated far-reaching health campaigns among their policy holders, with appreciable results in the direction of lengthening the span of human life.

Speculation. The fact that prices of particular goods and of goods in general tend constantly to rise and fall has already been called to the reader's attention (see Chapter XI). To the modern business man the risk of price change, with all its possibilities of loss, is one of the most serious of the hazards of enterprise. This is a hazard of which most entrepreneurs would gladly be relieved, and there has developed in consequence a device for passing on the risk of price changes. This is *speculation; i.e., the buying and selling of goods for the sake of the profit from price changes.* In many lines of business there are professional speculators, men who have sufficient confidence in their ability to foresee price changes to be

willing to assume the risks involved and to relieve those who prefer safety to risky profits. To be a successful speculator it is necessary to forecast the movement of prices, not absolutely, but simply better than the others in the market. Despite the fact that the existing price indicates the composite judgment of all buyers and sellers in the market, the speculator pits his opinion, manipulative skill, and luck against this judgment, and if he forecasts correctly he wins a larger return on his capital than presumably he can get for it in a line where his opinion is on a par with that of others. Faulty forecast means loss, which the speculator must bear.

The technique of speculation. The nature of the speculator's operations depends upon his judgment of the future price movement of the thing — commodity, land, government bond, corporation security, etc. — in which he is dealing. If the wheat speculator believes that the price of wheat will rise, he buys wheat now in order to hold it and sell it later at the higher price; he is a "bull," as the market jargon has it. If he foresees a fall of price, he sells now, expecting to buy later at the lower price; he is a "bear."

It is not necessary that the wheat be immediately accepted and held by the speculative buyer or that the speculative seller be actually in possession of wheat and make immediate delivery. Much speculative dealing is not for immediate physical transfer of the good, but for future delivery, called "dealing in futures." The "bull" on the wheat exchange, expecting a rise in price, buys wheat today to be delivered to him at a specified future date. When that date arrives, or before if he wishes, he sells this wheat for delivery on the day when it will be delivered to him. He is said to be "buying long." On the other hand the "bear," since he believes the price will fall, "sells short"; that is, agrees today to deliver wheat at a specified future date. He may not, probably does not, own today any wheat. On the specified date he will buy the wheat to deliver on his contract, or he may at any time before buy wheat to be delivered to him on the date when his selling contract matures.

Speculation equalizes prices. In any of the great speculative markets, such as the wheat market, the cotton market, the stock

exchange, etc., the speculators account for a considerable part of the demand and supply of the good in question. They have therefore a profound or even predominant influence in the determination of the price. And their influence is generally in the direction of stability of price, ironing out the extremes of price fluctuation. When the price is low, presumably because the demand is slight and the supply large, the effect of the bull's buying is to increase the demand and therefore to raise the present price or keep it from falling lower. At the later date, if he has judged correctly, the bull offers wheat for sale when the price is presumably higher, and the effect of this additional supply is to cause the price to fall or to keep it from rising higher. Raising the present low price and causing the future high price to fall therefore brings present and future prices more nearly into equilibrium. The bear's operation has a similar effect on prices. He sells today, when the price is high, thus adding to the supply and tending to lower the price. At the future date, when the price is lower, his buying tends to raise it. The whole class of speculators, thus inclining to buy when prices are low and sell when they are high, have a stabilizing effect upon price fluctuations.

This relation of cause and effect is simple enough when the dealing is for immediate delivery. In the case of dealings in futures the equalizing influence is not quite so obvious. We must remember that at the present time there are two prices, the price for immediate delivery, for example "cash wheat" as it is sometimes called, and the present price for wheat available, let us say, three months from today, or "future wheat." Each price is the result of its own separate demand and supply, yet the demand and supply for present and for future wheat are so closely related that there is a normal "spread" between the prices, which will be maintained over considerable periods of time. When the bear for example sells wheat for future delivery he enters upon the present market for futures on the supply side, and his action has the effect of depressing the price of future wheat. The "spread" therefore between the prices of present and future wheat becomes abnormal, and those interested in "cash" wheat are thus warned

that the speculators believe the price of wheat in the future will drop. Millers will feel that it is not so necessary to buy wheat now in order to be able to fill future contracts for flour, and they will diminish somewhat their present rate of buying; on the other hand those in possession of stocks of wheat will become somewhat uneasy lest they be left with unsold wheat on their hands, and they will be inclined to be less uncompromising in their attitude towards prices. The resultant decrease in demand and increase in supply of present wheat will cause the price to drop, and present and future prices will again be in alignment. When the "bear" is forced to fulfill his contracts he must enter the market to purchase "cash" wheat, and obviously he adds to the existing demand, thus tending to raise prices. The bear operation therefore tends to equalize prices by lowering the present high prices, and later (when the future has become the present) by raising the lower prices. Corresponding results follow the operations of the bull, who today buys long for future delivery and later sells the wheat which he has bought.

Speculation promotes economical consumption. As a corollary to the equalization of prices, speculation performs another important service. When goods are abundant and therefore cheap, the speculator's influence tends to take goods from the present market and hold them for sale at a time when there is relative scarcity. Present consumption will be somewhat curtailed in order that there may be more ample supplies for the future time. Without the operations of the speculator plenty and famine are likely to follow each other, but with more or less accurate forecasting of price changes the risk of famine is almost entirely avoided. In other words the speculator creates time utilities when he carries goods over to a period when the demand is more urgent. One of the earliest examples of speculative wheat dealing is seen in the dealings of Joseph in Egypt. During the years of plenty he accumulated stores of grain and other provisions to tide the Egyptians over the seven lean years, doubtless with profits proportionate to his shrewdness in applying his inside information as to future price movements. In the same way speculation

makes it possible to utilize present stocks more fully and thus avoid carrying goods over from a period of scarcity to a period of relative abundance.

Usefulness of the speculative market. Active speculation in any good, whether it be wheat or shares of stock, implies a market where price is determined by demand and supply with a degree of accuracy unknown in the non-speculative markets and where the process of price determination is continuous and uninterrupted. In such a market it is always possible to ascertain the exact price at which the good can be bought or sold, and anyone who wishes can always be certain of the ability either to buy or to sell at that price. In the absence of speculative dealings, it is doubtful if many commodities or securities could show sufficient exchanges to create this sort of a market. The presence of such a market is of inestimable utility to investors, manufacturers, merchants, contractors, and others outside the group of the speculators themselves. Compare the ease with which the administration of an estate can dispose of bonds and stocks listed on the exchange with the erroneous difficulty and material loss which frequently attends the sale of real estate or unlisted securities. It is the speculative stock exchange which enables bankers to regard their listed securities as a "secondary reserve," and it is the absence of such a market which bans real estate as a banking investment.

Hedging. In addition to its simple use as thus far considered, speculation offers a peculiar method of risk avoidance, known as *hedging*. Two illustrations of its use — by a miller and by a grain dealer — will serve to disclose its essential features.

Suppose a miller of Peoria, having noted the cost of materials, labor, and equipment for milling, contracts on June first to deliver to a jobber on July first 1,000 barrels of flour at ten dollars a barrel. The ten dollars are sufficient to cover the cots of milling and delivering the flour together with a fair profit, provided the price of wheat does not rise. Of course the miller might buy at once and hold the entire quantity of wheat required for making this flour. But the miller is seldom in a position to do this; he may not have adequate storage facilities, and he may not find in the

samples on June first the exact kind of wheat he needs. The former obstacle might be surmounted by taking advantage of simple speculation and buying wheat for delivery at future dates as required by his milling operation, but the difficulty of finding exactly the right samples would remain. For these and other reasons, millers generally prefer to buy their wheat in moderate quantities from time to time as they need it. But this involves the risk of loss from a rise in the price of wheat. In this situation, the miller has resort to *hedging, the essential feature of which is the elimination of risk by taking an equal position on both sides of the market.*

On the very day our miller contracts to deliver the flour he buys on the Chicago Board of Trade 3,000 bushels (by assumption enough to make the 1,000 barrels of flour) of July wheat. Of course he does not intend to grind the July wheat; it is not deliverable till during July, whereas the flour is due July first. He buys it because it affords him a hedge until he does find occasion to buy his milling wheat. He waits, let us say, until June sixteenth, when on the basis of samples he buys 3,000 bushels of milling wheat and instantly sells 3,000 bushels of July wheat. Now if, as is the usual case, cash wheat and July wheat differ by a normal spread, say four cents, when one rises or falls the other does likewise, staying four cents apart. Therefore if by chance the price of cash wheat rose 13 cents between June first and sixteenth, the miller's costs would increase by 13 cents a bushel above the figure he used in calculating the price at which he agreed to deliver the flour. But it is also plain that the July wheat also rose 13 cents above its cost by the time it was sold June sixteenth. Therefore the excess which he has to pay for his milling wheat he has gained on his July wheat. If the price had fallen he could have bought his milling wheat lower but would have lost on his July wheat a similar amount. In this manner, whether the market price rises or falls, his milling profit is protected.

Turning now to our second example, we may assume that an elevator at Wichita ships grain to Galveston directly, a distance requiring 20 days' time, but hedges on the Kansas City exchange.

Let us suppose that the normal difference between Wichita prices and Galveston prices is 15 cents — a difference sufficient to cover the cost of shipment from Wichita to Galveston plus the grain dealer's profit — and that prices at Kansas City are likewise higher, say by 8 cents, than the Wichita prices and for the same reason. Let us suppose also that on January first May wheat was higher than cash wheat by 5 cents, a difference caused by the cost of carrying. These being the conditions on January first, the Wichita elevator notes that on this day the price of cash wheat is $1.00 in Wichita and $1.15 in Galveston. With this difference, there is a fair profit in shipping wheat to Galveston. But, since the wheat cannot be sold there until it arrives, about the twentieth of January, there is the risk of a fall in price which might wipe out the profit or turn it into a loss.

To eliminate this hazard, the elevator hedges. It ships, let us say, 10,000 bushels of wheat to Galveston and immediately sells on the Kansas City market 10,000 bushels for May delivery at a price of $1.13. The grain arrives at Galveston on January 20 and wheat prices having in the meantime risen by eighteen cents, is sold for $1.33 a bushel. Immediately the Wichita elevator buys on the Kansas City market 10,000 bushels of May wheat to cover its short contract, paying $1.31 per bushel.

The price situation in this example may be better visualized with the aid of the following table.

	Wichita		*Galveston*		*Kansas City*	
	Jan. 1	*Jan. 20*	*Jan. 1*	*Jan. 20*	*Jan. 1*	*Jan. 20*
Cash wheat . .	1.00	1.18	1.15	1.33	1.08	1.26
May wheat . . .					1.13	1.31

The results of the hedge are as follows:

(*a*) On the May wheat:

Selling price, January 1	$1.13
Cost, January 20	1.31
Loss	.18

(*b*) On the wheat shipped:

Possible selling price at Wichita,	Jan. 1	$1.00
Actual selling price at Galveston,	Jan. 20	1.33
Gain		.33

Net gain from both transactions, $33 - 18 = 15¢$, which by assumption is the gross margin at which the Wichita elevator can handle Galveston shipments and make its regular trade profit, and which was the margin on which the elevator made its decision to ship on January first. If the price of grain had fallen during the time of transit it would have lost on the grain shipped but have gained on the May future.

The essential features of hedging are thus brought to light, especially these three: (1) no profits are made or are meant to be made from the hedge itself, (2) it is employed to protect a trade profit — the miller's for milling flour, the grain merchant's for dealing in grain — against loss occasioned by change of price, and (3) it depends upon the existence of a " normal spread " between the prices prevailing in two markets, the prices of two grades of the same commodity, the prices of convertible stocks and bonds, etc.; if this normal spread is not maintained the hedger may lose or gain by the amount the spread departs from normal.

Social service of speculation. Manufacturers, merchants, and entrepreneurs in general find in speculation a welcome relief from the hazards of price changes. The development of professional speculators is another example of the division of labor, which tends to increase the efficiency of our society. Relieved of some of the risks of business, the manufacturer or merchant can devote more time to the acquisition and employment of skill in his own specialty.

Moreover the transfer of risks from the relatively ignorant and uninformed to men who are experts in their particular field means a reduction in the hazard. Less risk attaches to a contract for the delivery of wheat at a given price at some future date made by an able speculator than to a similar contract made by a small local grain dealer, for it may be presumed that the speculator has the wider and more profound knowledge of world conditions affecting the demand and supply of wheat.

More than this, speculation, by equalizing prices, tends to do what insurance also does; namely, diminish the very events against which it operates. And finally there is the economy in consumption over long periods of time which speculation promotes. All these are services to society of great importance, well worth the price paid the speculating class in the form of speculative profits.

Investment, speculation, and gambling. That speculation is not an unmixed blessing to society is principally due to the fact that the machinery of speculation is available for the gambler as well as for the speculator.

It is not always easy to distinguish between the three concepts of investment, speculation, and gambling. The difference between speculation and investment lies very largely in the attitude of mind toward risk. The speculator is willing to assume the risks involved, whereas the investor prefers a moderate profit with the least possible risk. Inevitably he must assume some risk in any investment, but he reduces this to the minimum. The investor therefore tends to buy the securities of a stable government or the first mortgage bonds of well-established corporations or stocks of established reputation, and he looks for his gain primarily to the annual income of interest or dividends. The speculator on the contrary is attracted by the stocks that are likely to fluctuate in price and, thinking less of the income from dividends, buys or sells as he thinks the price is going up or down.

As we have seen, successful speculation is of advantage to society as a whole through the reduction in the volume of risks which it effects. Gambling, on the contrary, is marked by the creation of artificial and unnecessary risks. Since one gambler gains what another loses,[1] gamblers as a class cannot gain. Within the class, the especially shrewd and the dishonest gain at the expense of the others. The average outsider, since he is betting against experts and professionals, has little chance of success at gambling, whether on horse racing or on price movements in the pit,

[1] In any bet between men of equal incomes the presumption is that, measuring loss and gain in terms of utility rather than in terms of money, one man loses more than the other gains, for according to the law of decreasing utility the man who loses a bet of $100 loses more in utility than the man gains who wins the bet.

the stock exchange, or the realty market. The social verdict which considers the gambler a parasite on society is probably fully justified.

Finally, quite aside from its effect on the gamblers themselves, gambling is a source of disturbance in the market and so a cause of loss to society. The otherwise intelligent citizen who decides to take a " flyer " in wheat or securities in the hope of securing " easy" money is more likely than not to be on the wrong side of the market. If he buys wheat in the expectation of a rise and the price actually falls, not only does he lose, but his operation accentuates the price difference between the future and the present, since his purchasing in the present will raise the present price and his later selling will still further depress the low price of the future. Of course even the experienced speculators do not forecast price movements accurately at all times, but any speculator who is habitually amiss in his forecasts soon loses the power to speculate.

The corner. It occasionally happens that a few powerful speculators can secure sufficient control of the supply of some commodity to enable them to determine the price. This is one of the real evils of speculation, but it is one whose importance is greatly overestimated by the public. Successful corners have been very rare and have lasted only a few days at the most. Price control through corners has little effect on the public, as the corner is usually aimed at other speculators who have made contracts for future deliveries and who must buy the commodity whatever the price in order to fulfill their contracts. The occasional spectacular winnings of those who have engineered successful corners are obtained from the other speculators and are not a cause of loss to the general public, except as its members have voluntarily taken gambling chances on the market.

Judgment of speculation. Some of the most favored and persistent of the economic fallacies which popular opinion cherishes centre about speculation and the speculative market. These fallacies generally rest upon ignorance of the laws of price, though they are to some extent excusable on account of the difficulty, already noted, of distinguishing always between investment,

speculation, and gambling. When it is charged that a great part of the dealing on the speculative exchanges is gambling, the truth of the charge must be conceded. However the seriousness of this evil, from the public point of view, is generally much exaggerated.

The spectacular profits sometimes acquired by successful speculators have had a dazzling effect and have fostered the belief that the speculators as a class are amassing huge fortunes out of the producers on the one side and the consumers on the other. As a matter of fact, extraordinary speculative profits are generally at the expense of equal losses by other speculators, and it is not at all probable that the net income of the whole class of speculators is more than a fair reward for the useful services which we have seen that they perform.

The most persistent charge is that on the exchanges shrewd and unscrupulous speculators are able to fix prices at their own sweet will, in violation of the law of demand and supply and to the enormous injury of both producers and consumers. Waiving the difficulty of understanding how the same price can injure both producers and consumers, the fact is that nowhere in the world is there so close approach to determination of price by supply and demand according to economic law as in the great speculative exchanges. It is here that we must always turn for our practical illustrations of the theory of price. Except in the rare and brief case of a corner, it is quite impossible for the speculators to "rig" the price, for the simple reason that two groups, with opposed interests, are always working against each other. The speculative market offers the minimum of friction to the operation of the laws of demand and supply.

The opponents of speculation often go to the length of demanding restrictive legislation or regulation, the prohibition of "futures," or even the entire abolition of the speculative exchanges. In view of what we now know of the services of speculation, such action would be a social calamity of the first magnitude.

An interesting illustration of these truths occurred at the time of the Civil War. "When the greenbacks became the regular money of the country, there was established in the city of New

York a 'gold exchange,' a speculative market where people met daily to buy and sell gold, and where the price of gold for the whole country was determined. As the greenbacks depreciated, there were great complaints about the high price of gold, and many people charged that the speculation on the gold exchange was responsible. This was exactly the same sort of complaint that we hear from time to time now about the New York Cotton Exchange or the wheat speculation on the Chicago Board of Trade. Congress gave ear to these complaints and passed a law abolishing the gold exchange and forbidding all speculative dealings in gold. The result was immediate and disastrous. Nobody now knew where he could get gold or sell gold or what the price was. There was now no common point where all the supply and all the demand could come together and determine the price. Instead there was bargaining and higgling at various places, and widely differing prices were quoted in different places. The price, instead of falling as hoped, immediately rose higher than ever. It was 198 on the day the law was passed. The next day it was 208; the next day, 230; and in less than two weeks it had risen to 250. This was the natural result of the bidding of people who had to have gold and who no longer knew where to get it or what was a fair price. The demand for the restoration of the exchange was so prompt and so strong that Congress repealed the law just two weeks after its passage. From that time on the gold exchange did business till the greenbacks came to par in December, 1878." [1]

An impartial weighing of its merits and demerits can leave no question as to the net advantage of speculation to society as a whole. There are certain aspects of speculation which may justify some public regulation, but any action which would seriously interfere with the normal operation of speculation would be a costly blunder.

[1] Fairchild, *Essentials of Economics*, p. 370.

Suggestions for Further Reading on Part IV. Problems of the Business Organization

I. Railway Transportation

Jones, E. *Principles of Railway Transportation.* New York, 1924
Acworth, Sir W. M. *The Elements of Railway Economics.* Revised edition. Oxford, 1924
Vanderblue, H. B. and Burgess, K. F. *Railroads: Rates-Service-Management.* New York, 1923
Hadley, A. T. *Railroad Transportation: Its History and Its Laws.* New York, 1897
Raper, C. L. *Railway Transportation.* New York, 1912
MacGill, C. E. (and others) *History of Transportation in the United States before 1860.* Washington, 1917
Brown, H. G. *Transportation Rates and Their Regulation.* New York, 1916
Cunningham, A. J. *American Railroads: Governmental Control and Reconstruction Policies.* Chicago, 1922
Dixon, F. H. *Railroads and the Government.* New York, 1922
Haines, H. S. *Problems in Railway Regulation.* New York, 1911
Haney, L. H. *The Business of Railway Transportation.* New York, 1924
Johnson, E. R. and Van Metre, T. W. *Principles of Railroad Transportation.* New York, 1916
Sharpman, I. L. *The American Railroad Problem.* New York, 1921
Ripley, W. Z. *Railway Problems.* Boston, 1913
——— *Railroads: Rates and Regulation.* New York, 1912
——— *Railroads: Finance and Organization.* New York, 1915

II. Ocean Transportation

Calvin, H. C. and Stuart, E. G. *The Merchant Shipping Industry.* New York, 1925
Johnson, E. R. and Huebner, G. G. *Principles of Ocean Transportation.* New York, 1919
Hough, O. B. *Ocean Traffic and Trade.* Chicago, 1914
Marvin, W. L. *The American Merchant Marine, Its History and Romance from 1620 to 1902.* New York, 1902
Bates, W. W. *The American Marine.* New York, 1902

III. Industrial Monopoly and its Control

Jones, E. *The Trust Problem in the United States.* New York, 1921
Jenks, J. W. and Clark, W. E. *The Trust Problem.* New York, 1917
Durand, E. D. *The Trust Problem.* Cambridge, 1915

HANEY, L. H. *Business Organization and Combination.* New York, 1913
HENDERSON, G. C. *The Federal Trade Commission.* New Haven, 1925
MEADE, E. S. *Trust Finance.* New York, 1903
RIPLEY, W. Z. (Editor) *Trusts, Pools, and Corporations.* Boston, 1905
STEVENS, W. S. (Editor) *Industrial Combinations and Trusts.* New York, 1913
United States Commissioner of Corporations — Special Reports on the Beef Industry, Steel Industry, International Harvester Company, The Standard Oil Company, etc.
United States Federal Trade Commission — Special Reports on Various Industries
WYMAN, B. *The Control of the Market: A Legal Solution of the Trust Problem.* New York, 1911

IV. RISK, INSURANCE, AND SPECULATION

HUEBNER, S. S. *Life Insurance.* New York, 1915
———— *Property Insurance.* New York, 1920
GEPHART, W. F. *Principles of Insurance (Fire and Life),* 2 volumes. New York, 1917–1920
WINTER, W. D. *Marine Insurance.* New York, 1919
HUEBNER, S. S. *Marine Insurance.* New York, 1922
BRACE, H. H. *Value of Organized Speculation.* Boston, 1913
MEEKER, J. E. *The Work of the Stock Exchange.* New York, 1922
HUEBNER, S. S. *The Stock Market.* New York, 1922
BOYLE, J. E. *Speculation and the Chicago Board of Trade.* New York, 1920
NOURSE, E. G. *The Chicago Produce Exchange.* Boston, 1918
HUBBARD, W. H. *Cotton and the Cotton Market.* New York, 1923
EMERY, H. C. *Speculation in the Stock and Produce Exchanges of the United States.* New York, 1896

PART V

THE DISTRIBUTION OF WEALTH AND INCOME

CHAPTER XXXIII

ECONOMIC RENT IN AGRICULTURE

The problem of distribution. At various points in our study of Economics we have observed that man's economic activity centres around the question of income. Thus far we have been concerned with two phases of this subject: the problem of income from the standpoint of *production;* and the same problem from the standpoint of *exchange.* If we proceed from either of these starting points and carry our investigation far enough, it will eventually embrace the whole of society's economic structure.

It will now be our purpose to take a different point of view toward the problem of man's economic activity and lay emphasis upon a third group of problems connected with it. Society must settle in some way the all-important question of sharing or dividing the product of industry among its members. This process of parcelling out the product of coöperative economic activity is called *distribution.* Distribution, as the term is used in Economics, is concerned with those forces which determine the relative and absolute sizes of the incomes of various individuals. It is a study which may be pursued with two different objects in view. First we may attempt to measure the actual incomes of the different individuals and to explain the inequalities which appear among them. This we might call a study of *personal distribution.* Obviously it would require the collection and analysis of a mass of statistics dealing with the incomes of society's members. Or secondly we may classify incomes on the basis of their sources, or of the services rendered by their recipients in exchange for them. This would lead us into a study of the different factors of production and of the economic laws which control their remuneration. Such an inquiry would be called a study of *functional distribution.*

We shall study both of these phases of the problem of distribution in the following chapters, postponing our study of personal distribution until we have made a preliminary study of the functional types of income. There are four of these functional types: economic rent, interest, wages, and profits. Each of these types of income involves certain peculiar problems and is determined by independent forces.

The technical terms of distribution. As we follow our study of distribution further, we shall find that the commonplace names used to denote the different types of income are not the clear and definite concepts which on first sight they appear to be. The terms "rent," "interest," "profit," are so often used in different senses by the economist, the business man, and the general public, that it is necessary at the outset of their study to define them with some care. In common usage the term "rent" denotes any payment made by one man to another for the use of an article of wealth, when the payment is proportioned to the *physical unit* of wealth rather than to its value. Thus it is said that men "rent" pianos, automobiles, plots of ground and buildings, or masquerade outfits, whenever they discharge their obligations to the owner of these things by paying him so much for the use of each physical unit. These expressions are all in good usage, and everyone knows what they mean.

Economic rent an income from land. But as used in economic science to denote a peculiar sort of income, the term *economic rent* has a much more restricted meaning. It is that part of the total income of society which accrues to the landowner; not, indeed, as a reward for any productive service performed by the owner himself, but from the use of his land for productive purposes. It is one of the consequences of our institution of private property that land owned by the members of society can be put to productive uses only with the owners' consent and upon such terms as they will accept. If the economic forces working in society at a given time bestow on a certain piece of land power to contribute a definite, ascertainable quantity of income to the total social fund, over and above what could be produced by society's capital and

labor without the use of this land, its owner can demand that this quantity of income (or its value) be given to him as a condition for the use of his land.

He has two ways of enforcing this demand. By holding his land idle, he can force others to bid for the privilege of using it in some business enterprise until the highest bidder offers him a money payment approximately equivalent to the income-producing power of the land. Or he can work his plot of ground himself and thus derive directly the benefit of its productive power. In the first case, the owner of the land will receive a definite money payment, distinct from all his earnings as a worker or investor, which should be called *contractual or explicit economic rent.* In the second case also he will receive an income, of approximately the same value; but it will be merged in a composite flow of benefits representing the total result of his activities as worker, investor, and land owner. When mixed in this fashion in the total income which accrues directly to the owner, the portion of income traceable to the productive power of land should be called *implicit economic rent.*

Definition of economic rent. Certain qualifications should be added to make clear beyond question what this economic concept implies. In the first place, it must be remembered that, as was pointed out in Chapter II, "land" is a generic name given in economics to all the resources of nature afforded by man's physical environment, including the mineral deposits of the earth, the forests, the water power, the medicinal springs, the fisheries, etc. These physical properties of the earth are in nature different from those articles of wealth in the making of which man himself has had a hand. This difference is of peculiar importance in that part of man's economic activity which has to do with the distribution of income, as will appear in the sequel.

In the second place, a further word of caution must be added with regard to those cases where economic rent takes the form of a contractual or explicit payment. If we examine the ordinary case in which land (or other natural resource) is leased by its owner, we shall find that the lease covers articles of wealth imbedded in the land which are the result of man's labor and not the gift of

nature. The tenant in the city, for example, commonly obtains the use of buildings, sidewalks, and other improvements in addition to the use of the site. In the country districts, fences, barns, drainage systems, orchards, go with the land. In both cases a single payment is made for the use of this composite unit of land and capital. This payment, from the standpoint of economic science, is made up of two very different elements: *interest on capital* and *economic rent of land.* In our study of economic rent we are concerned only with that part of the landowner's income which may be attributed to the use of the land itself unimproved and devoid of capital.

In the light of these explanations, we are now prepared to define the concept we are using. *Economic rent is a net income from the use of land, distinct from all income from labor or the use of capital.* To avoid ambiguity, we will employ the full term *economic rent* when referring to this peculiar type of income. When required to refer to a contract payment which includes other things than economic rent, we will use the word *rental.*

Economic rent equated to the productivity of land. We have said that this income in any given case measures the contribution which the land has made to the total income fund of society. But how is it possible to measure or evaluate this contribution? It may be said with truth that the entire social income is attributable to land (and other natural resources) in the sense that without the coöperation of this indispensable agent man could produce nothing at all. With equal verity the counter statement can be made that all social income is due to labor, since nature, even in its most bountiful moods, cannot satisfy men's wants without the expenditure of at least some slight effort on their part. But in the productive system of modern society we are not compelled to come to terms with all the owners of land at one stroke; land is not employed as a single mass, but piecemeal in certain definite quantities — acres or square feet. Payments are not made to landowners-in-general for the use of land-in-general, but to separate individuals for the use of definite plots. If all owners of natural resources were to combine their interests and withhold from us the entire

surface of the earth, they could make us pay tribute to the full extent of the social income — allowing only for such starvation wages as would keep the rest of us at work. There have been cases in history which approximated this condition of affairs. The despotic sovereign of a vast domain has sometimes exacted from his enslaved subjects an income which approached this limit. This income, however, was not economic rent, but a tribute or forced levy. In our competitive order, owners of land make their bargains as individuals; and no one of them, by withdrawing his land from use, can reduce the social income to zero. At the worst, he can only force us to resort to plots of ground or other natural resources inferior to his own and can demonstrate the fact that society's labor and capital employed without the coöperation of his particular portion of the earth's surface will produce a smaller income than we would have if we came to terms with him.

If this matter is still puzzling to the student, it can be made clear by recalling that the same principles govern the exchange value of all goods bought and sold in the market. Foodstuffs, all will agree, are vital to the existence of mankind; yet we do not barter our lives or all our possessions with the suppliers of these things. The reason, as we discovered in our study of value, is because we value them in units and not in the general mass and pay that price for each of the units of a homogeneous quantity which measures the contribution which one unit makes to our enjoyment. So it is with land. All acres of land do not command the same economic rent, because they differ in productive efficiency by reason of differences in fertility, in location, in climate, or otherwise. But the part played by any one acre in the production of income may be discovered by withdrawing this acre from use and thus compelling our capital and labor to have recourse to the remaining portion of the land area.

The principle of diminishing returns. But why does the withdrawal of a given plot of ground from use cause the social income to fall? Why cannot our labor and capital be applied to what is left of the land area without causing a diminution of income? To speak concretely, suppose we have to do with a tiny community

whose entire land area consists of a hundred acres. If one, or ten, or fifty of these acres are withdrawn from use by their owners, could not the members of the community betake themselves with their tools to what is left and produce as much as before? It is apparent that if this could be done, the owner of no single acre could compel anyone to pay him an income for the privilege of using it. For to the other members of the community it would be immaterial whether this particular acre were used or allowed to remain idle; their total income would be the same in any case. Accordingly, they would make no bargain with the owner giving him a share in the total income as a landowner in addition to his reward as a worker or supplier of tools. But in general this supposition is unsound; neither in the little community we have pictured nor in the great society in which we live can labor and capital be withdrawn from a part of the land and concentrated on the rest without causing a diminution in the total income. The reason is that the exploitation of land and other resources is governed by the *law of diminishing returns*. It is because of the existence of this law of nature that economic rent itself exists as a separate type of income.

Diminishing returns in agriculture. Reference to the principle of diminishing returns has been made in connection with other parts of our study. An inquiry into its relation to the problem of economic rent will be simplified if we limit ourselves, in the first instance, to the case of agricultural lands. The income from this sort of land usually appears in tangible form which permits of objective study of the effect of diminishing returns. Accordingly we shall confine our study in this chapter and the next to lands used in agriculture and proceed in the chapters which follow to apply our conclusions to urban lands and to mines and other resources.

As applied to agricultural lands the principle may be stated as follows: *after a certain point has been reached in the cultivation of a given area of land, additional applications of labor and capital to that area, though they will increase the total product, will not increase it in proportion to the increase of capital and labor employed.*

This can best be explained by means of an illustration. Let us take, for example, a plot of land, twenty-five acres in area, which is used for the production of oats. Assume at the outset that the entire area is tilled by the proprietor with his complement of seed, tools, and other capital; and that it produces under such tillage 500 bushels of product. Because of the economy of time and effort made possible by division of labor, a second worker with an equal equipment of capital can, we may assume, double the product, making the total output 1,000 bushels. But this doubling of the product by doubling the effort expended cannot go on indefinitely; the time will come when an additional unit of capital and labor will not increase the output proportionately. Let us suppose that this point is reached when a third worker and his equipment are employed in cultivating the twenty-five-acre farm. Whereas each of the first two workers with his equipment added 500 bushels to the output, the third man similarly equipped adds 400, thus making the total product 1,400 bushels. A fourth man with the same equipment adds 300 bushels, making the total 1,700. And a fifth man adds 200 bushels, making the total output of the farm, when cultivated by five men, 1,900 bushels of oats. It will be seen upon examining this illustration that, although the total product of the land continues to increase with the increase of labor and capital applied to it, the *marginal* product — that is, the amount attributable to the last man and his capital equipment — decreases as the third, fourth, and fifth man are added. This declining marginal product is the symbol of diminishing returns. It is characteristic of all extractive industries. To be sure, the point at which diminishing returns will appear is not the same for all grades of land, for all sorts of crops, and for all kinds of agricultural organization; but sooner or later the point is reached where nature responds to increased intensiveness of cultivation with a diminishing marginal product.

What is implied in diminishing returns. So important is this principle of diminishing returns in our study of economic rent — and indeed with regard to the other shares in the social income as well — that we may well pause to make clear its implications.

It should be noted, in the first place, that the statement of the existence of diminishing returns does not imply that small scale farming is more economical than large scale. No formulation of the law of diminishing returns can be made except on the assumption that we deal with a given quantity of land, which quantity is held constant while the expenditure of labor and capital is increased. When we compare the relative economy of small and large scale farming, we have to do with areas of different sizes. On neither of these areas, regardless of the number of acres it contains, can the labor forces be expanded indefinitely without eventually reaching the point where the product increases less rapidly than the increase of labor and capital expended. The principle of diminishing returns applies to land areas of every conceivable size — from the scrap of ground worked by the truck gardener to the great domain of the rancher. The areas of various sizes differ from each other in respect of the operation of diminishing returns only with regard to the point at which the marginal product will begin to fall. It may be true that a second man added to the cultivation of the truck garden might fail to double the product, while on the large ranch the number of workers might be increased to ten or twenty before diminishing returns became apparent.

In the second place, the principle of diminishing returns must not be confused with the tendency of the soil to lose its fertility after successive years of cultivation. The declining marginal product of the land which registers the operation of diminishing returns is not to be attributed to the exhaustion of the soil. It holds true in the first year of cultivation of virgin land as much as in the results obtained after many years of tillage. In stating the principle of diminishing returns we assume that we are dealing with a single season of cultivation; that, to refer again to our illustration, in any given year the twenty-five-acre farm would produce 500 bushels if tilled by one man, 1,000 bushels if tilled by two, 1,400 if tilled by three, etc.

In the third place, in conformity with this implication of a given point in time, the principle of diminishing returns rests on the

assumption that the technique of cultivation remains unchanged when successive units of capital and labor are applied to a constant area of land. The fact that improvements in agricultural machinery, fertilizing methods, and the science of tillage, have brought it to pass that the output per worker is larger now than formerly in no wise conflicts with the principle of diminishing returns. On each stage of the arts, from the beginnings of the rudest tillage to the modern stage of scientific agriculture, it has always been true that the marginal product of the land would have decreased if the number of workers per acre had expanded beyond a certain point. It is true that there have been many societies in the course of history in which, because of the relative scarcity of men in proportion to the land area, it has not been necessary to concentrate effort on a limited portion of the soil. Such societies may have passed many generations of their existence without experiencing the operation of diminishing returns in the exploitation of their natural environments. But the principle has held good potentially at all stages of their development.

Validity of the law of diminishing returns. When understood in this strict sense, as applying to a given area of land, a given period of time, and a given stage of the arts, the truth of the law of diminishing returns is supported by incontrovertible facts of experience. We have but to witness the action of any practical farmer to be convinced that he accepts its truth, however unwittingly. No farmer assumes that it is profitable for him in any given year to apply labor and capital without limit to the cultivation of his fields. After hiring a certain number of workers at a standard rate of wages and equipping them with capital at a fixed cost per man, he knows that further additions to his labor force "will not pay"; that is to say, additional men will not increase his output enough to make their employment worth while. But why should this ever be true except that increasing intensiveness of cultivation results in a decreasing marginal output? Or, looking at the matter in a broader social light, why are soils of differing degrees of fertility cultivated to the production of the same crop at the same time? Why not concentrate effort upon the very best

acre of wheat land and feed the world from this single source? If the labor of the first man applied to this most fertile acre results in a larger product than the same labor applied to inferior soil, why should not the same be true of the labor of the thousandth or the millionth man? Reduced to extreme cases, the existence of the law of diminishing returns is certainly self-evident; and it may be that we are laboring a point which is already sufficiently clear. But so essential is comprehension of this principle to an understanding of the theory of distribution in all its branches, that we must take time to make it clear even at the risk of repetition. If doubts as to the meaning of the principle still linger in the student's mind, he will be well advised to go over again the explanation of the last few pages, testing the law of diminishing returns by applying it to events of his own experience, before proceeding with his study of economic rent.

Diminishing returns translated into increasing marginal cost. We may now return to our illustration in order to subject it to further analysis. This analysis will be facilitated if we throw the illustration into tabular form, as follows:

GIVEN: A FARM OF TWENTY-FIVE ACRES

Number of workers equally supplied with capital	*Output of the Farm*	
	Total product in bushels	*Marginal product in bushels*
1	500	500
2	1,000	500
3	1,400	400
4	1,700	300
5	1,900	200

The third column of this table is significant for our purposes. It records the marginal products of the workers; that is, the part of the total product which is attributable to the presence in the labor force of the last man who joins it. When one man works alone, the total and marginal products are the same. When the labor force is expanded by the addition of another man equally equipped with tools, the total product rises from 500 to 1,000 bushels; hence the marginal product is still 500 bushels. Diminishing returns set in with the employment of the third man. A three-man labor force

with a proportional increase of capital increases the total product from 1,000 to 1,400 bushels; hence the marginal product of three men is 400 bushels. Four men with their tools produce 300 bushels more than three men; this is the marginal product of a four-man labor force. Similarly, it is shown that a five-man labor force produces a marginal product of 200 bushels.

It will be recalled that we have assumed each man in the working force to be equally well equipped with capital; on this assumption there is a uniform cost involved in employing each of the workers. If now we resolve this concept of an expanding working force into terms of monetary cost, thus presenting it in the guise in which it enters into the farmer's reckoning, we shall see that the principle of diminishing returns is but another expression of the *law of increasing marginal cost,* with which we have become familiar in our study of value. In doing this we must be careful to include all the costs of cultivation. First a compensation must be allowed for the labor and management of the proprietor, sufficient to induce him to put forth his efforts as a worker and entrepreneur, since this is certainly one of the costs of the undertaking. Secondly we must include among the costs a normal interest return on the money invested in tools and equipment. Finally his expenditures for labor, seed, and other materials must be added. In the following table we have simplified the matter somewhat artificially by assuming that all of these costs can be proportioned to the number of workers employed in cultivating the farm.[1]

STATEMENT OF INCREASING MARGINAL COST

Expenditure in dollars	*Total product in bushels*	*Marginal product in bushels*	*Average cost per bushel*	*Marginal cost per bushel*
$100	500	500	$.20	$.20
200	1,000	500	.20	.20
300	1,400	400	.214	.25
400	1,700	300	.235	.33⅓
500	1,900	200	.263	.50

[1] The chief element of artificiality in our mathematical illustration obviously consists in the elimination of all costs in the nature of fixed charges and overhead expenses. (See Chapter XVII.) Recognition of these costs, while complicating our analysis, would not in any way affect the conclusion. And, as a matter of fact, in a simple enterprise such as this, fixed and overhead costs are of slight practical importance.

Since, according to our assumption, the figures in the first column of this table represent all the costs of producing oats on this farm of twenty-five acres, it is apparent that if the various quantities of produce are sold for enough to return these costs there will be sufficient inducement to the proprietor to continue the operation of the farm. The second and third columns of the table, showing the total and marginal products, are the same as in the preceding table and require no explanation here.

It is with the figures appearing in the fourth and fifth columns that we are at present concerned. These figures show the average and marginal costs of the product. The table of average costs per bushel is obtained by dividing the total product (second column) into the total expenditure (first column). It will be noted that the average cost rises after the point of diminishing returns has been passed. In the final column is recorded the marginal cost of production. These figures are obtained by dividing the marginal products of the working force into that part of the total cost which was involved in the production of the marginal product. The first marginal product is 500 bushels; this product involves a cost of $100; the marginal cost is, therefore, twenty cents. The marginal product of two men is also 500, produced by an additional expenditure of $100; hence the marginal cost remains at twenty cents. The third marginal product, added by the expenditure of an additional $100, is 400 bushels; the cost of this marginal product is accordingly twenty-five cents per bushel. By extending this computation we find that the cost per bushel of the fourth and fifth marginal products is thirty-three and a third and fifty cents, respectively. We note again that the marginal cost of production rises after cultivation is pressed beyond the point of diminishing returns. Clearly, therefore, diminishing returns is but another name for the familiar concept of increasing marginal cost — the same concept presented from another point of view.

Compare now the figures of average cost with those of marginal cost. At the outset, these two costs are the same; and they would continue to be the same indefinitely but for the operation of diminishing returns. But with the beginning of diminishing

returns the marginal cost rises faster than the average cost, thus causing a gap or "spread" to appear between these two costs, which grows wider as a larger and larger product is forced from the soil. Referring again to our table, we observe that when the output of the farm is 1,400 bushels, marginal cost is higher than average cost by a difference of 3.6 cents per bushel. When the total product is 1,700 bushels, this difference between the two costs becomes 9.8 cents; and when the product rises to 1,900 bushels, the difference increases to 23.7 cents. The fact that these two costs of production differ in extractive industries has already been made familiar to us through our study of value under conditions of increasing cost. If successive units of a given good cost more and more as the supply is increased, it stands to reason that the cost of the last unit must be greater than the average cost of all of them. Our illustration merely gives a certain definiteness to this familiar truth.

Marginal cost equated to market price. In making this detailed analysis of a hypothetical example, we may seem to have lost sight of our main task — the problem of economic rent. But the bearing of this discussion on the problem we are studying will become clear when we turn from a consideration of the costs of our farmer to take note of his receipts or income. Since he is producing goods for sale on the market, his money income will be determined by the price set upon his product through the operation of the forces of demand and supply. Now we have learned from our study of market price that the price of any good produced under conditions of increasing cost will tend to equal the marginal cost of production. If the price should fall below this point, all producers would suffer loss from the production of that portion of their output whose cost stood above the market price; and their reaction would be to restrict their output within such bounds as were necessary to avoid this loss. The restriction of supply would continue until the marginal cost of each producer was no higher than the price.

Our farmer, for example, will produce 1,900 bushels of oats only on the supposition that the selling price is fifty cents per bushel; or 1,700 bushels, only if the price is thirty-three and one third cents

per bushel; or any other quantity, only if the market price is high enough to cover the marginal cost of that quantity. This statement in fact sums up the relationship of each individual producer to the market. Each assumes that the market price, at any given moment, is beyond his control and makes it his concern to avoid producing any part of his output at a loss. When the price rises, he increases the application of labor and capital to the soil and thus advances the marginal cost toward equality with the new price, being induced to do so by the prospect of larger gains to be made from his increased output. Conversely, when the price falls he adopts the contrary course and contracts his production until he ceases to lose on the marginal units of his supply. These too are corollaries of the laws of value familiar to all students of the market.

Producer's surplus drawn from the difference between average and marginal costs. But all units of a given commodity sell in the market at the same price. Since this price is set by the marginal cost of production, which, after the point of diminishing returns has been passed, is higher than the average cost, it follows that the farmer receives in exchange for his product money income which contains a surplus profit. The extent of this surplus profit, for any given amount of product, may be reckoned in either of two ways. Assuming, for example, that the market price for oats is thirty-three and a third cents per bushel, we know that our farmer will produce 1,700 bushels. Selling this quantity at the price assumed, he receives a gross money income of $566.66. Reference to the table shows that his total cost is $400; hence he receives a surplus profit of $166.66. The same result can be reached by another method of computation. The difference between marginal and average costs for an output of 1,700 bushels is 9.8 cents. This is the farmer's surplus profit on each bushel sold; multiplied by the total quantity, 1,700 bushels, it gives a total surplus of $166.66. The second method of computation is, for purposes of illustration, to be preferred to the first, since it discloses the reasons for the existence of surplus profit — namely, the fact that marginal cost, which governs selling price, has risen above average cost, which determines the total expenditure of the producer.

The point should be emphasized that this part of the farmer's income is a *true surplus* over and above all payments required to induce production of the quantity offered for sale. It will be recalled that we embraced in the total cost of production interest on all capital used, wages for all hired labor, replacement of all materials consumed, and a remuneration for the labor and management of the farmer adequate to call forth his effort. Since it is important to avoid the appearance of having begged the question by a juggling of the conditions of our illustration, we should point out that no reasonable assumptions regarding this normal profit or reward of the farmer will change the results of our analysis. We may take any figure we please as a measure of this factor in the cost of production; it must still appear in our column of total costs. Changing our specifications of the remuneration to the farmer will alter our figures for average and marginal costs of production, but it will not alter the essential fact; that, after the appearance of diminishing returns, marginal costs will mount higher than average costs. If the farmer's product sold for no more than enough to cover his total cost (including the normal reward for himself), it would nevertheless prove sufficient to justify his operation of the farm. The fact that it sells for more than this is to be attributed to the laws of nature which cause our exploitation of natural resources to proceed under the principle of diminishing returns and to the common sense of producers which impels them to grasp all the gain that is open to them and to avoid loss on any portion of their output when the remedy lies ready to their hand. The surplus profit we have been describing is economic rent. A synonymous term for it, and one which fits exactly into the conditions of our illustration though it is less definitive of contract payments, is *producer's surplus*. We see that it is inevitably involved in the income derived from the cultivation of the soil after production has been pressed beyond the point where diminishing returns set in.

Contractual economic rent equivalent to producer's surplus. Thus far we have assumed that the owner of the land is also one of the workers engaged in cultivating it, as well as the manager of

the enterprise. The economic rent which accrues to him is an implicit income, not received as a separate and distinguishable payment but mingled with his total receipts from the sale of his product. He gains this portion of his income, not by virtue of his services as a workingman nor as a supplier of capital nor as a manager of industry, but solely by reason of his ownership of land. Private property in land, in a society which has pressed cultivation beyond the point of diminishing returns, places the owner in a position of advantage as compared with the other members of society. It affords him opportunity to live in greater comfort or to accumulate wealth faster than the landless man of equal industry and skill. Accordingly it is a position which landless men will covet and to attain which they will gladly assume financial costs.

Suppose, now, that our farmer wishes to retire from active labor and to that end offers his land for lease. He will of course expect that his earnings as a worker will stop. We may assume also that he turns his supply of agricultural capital — tools, machines, etc. — into cash and invests the proceeds in good bonds, thus continuing to receive that portion of his former income which consisted in interest. Obviously then his real financial sacrifice upon retiring from the land will be measured by the economic rent which formerly accrued to him. This accordingly will be the price he asks from prospective tenants for the use of his land. It is also a price which the tenant will be willing to pay; for, after subtracting this much from his gross income from the sale of his produce, he will have left enough to pay all the costs of producing and retain for himself a normal reward for labor and management. More than this he could not afford to pay, since that would make his lot as an independent farmer less attractive than as a worker in the employ of someone else. Less, the competition of other prospective tenants will not permit him to pay, since they too will place the same appraisal upon the advantage to be derived from the use of this particular plot of ground. Consequently the contract economic rent will tend to equal the surplus profit implicit in the cultivation of the soil by the owner. After the terms of the lease have been concluded we will then find the owner of land in an

unusual position as a sharer in the social income; with no further expenditure of labor and rendering no service as a supplier of capital (or, as the case may be, with adequate compensation for these functions) he will continue to draw an income from his privileged position as a landowner.

In our effort to simplify and make clear the major forces which govern economic rent as a social income, we have, it must be admitted, given the matter an appearance of precision and exactness which one does not find in the world of affairs. This qualification applies no less to the statement that contract and implicit economic rent equal each other than to other parts of our analysis. We can easily imagine cases of tenants who are willing to pay a little more than the economic rent in order to win a position of independence, or of landlords who consent to take a little less in preference to making extreme use of their bargaining power. Neither party to the lease contract can calculate the economic rent to a nicety, and when, as is normally the case, a certain amount of improvements goes with the land, it is beyond the power of human ingenuity to distinguish rent from interest to the last penny. Considerations of the probable future behavior of prices or of costs of production — considerations left out of account in our illustration because of our supposition of a fixed point in time — may also influence the calculations of tenant or landlord; though if these considerations affect the rental it will be because they indicate a change in the economic rent for coming years. Detailed qualifications of our statement may be multiplied. But they do not alter the essential principle after all is said. Competition among landlords on the one hand and among tenants on the other works continuously to the end that contract economic rents approximate in value the producer's surplus inevitably resulting from exploitation of land under conditions of diminishing returns. Only when lands are leased for long periods at fixed rentals, so that the changing conditions of the market upset the calculations of landlord and tenant, will the tenant be enabled to retain any considerable portion of the economic rent for himself or the landlord succeed in grasping more than this economic rent.

CHAPTER XXXIV

ECONOMIC RENT IN AGRICULTURE (Continued)

Cultivation of lands of different productivity. In our discussion thus far, we have dealt with a single piece of land of uniform quality per acre under cultivation by a single proprietor or tenant. If all the land of the earth were of one quality, rent would arise upon it as soon as the amount of product demanded by consumers had grown so large that a portion of the supply must be produced under conditions of rising cost. But the arable lands of the world are not of uniform grade; differences in productivity exist because of differences in inherent fertility, in proximity to market, in facilities for acquiring the necessary supplies of capital and labor, and so on. These differences in quality, in this day of commercial agriculture, appear to the cultivator in the guise of varying costs of production on different plots of ground. The typical farmer in the United States usually has within the limits of his farm many plots of varying fertility, ranging from the choicest fields, where the product per unit of expenditure is high, to patches of barren soil that are allowed to lie waste. Similar gradations appear as between the holdings of different farmers. For each different crop there are lands whose nearness to market and abundant fertility afford the largest output with the least expenditure of effort, others less well situated or less fertile where the average cost of production is higher, still others far removed from the market or composed of thin or impoverished soils whose cultivation at existing prices of farm produce does not pay. We may bring our discussion of the laws of rent into closer harmony with these facts of experience and incidentally permit of further analysis of the social bearings of this type of income, if we pause now to take account of differences in the income-bearing powers of land.

Let us suppose, to make the case simple, that there are three different grades of soil all devoted to the cultivation of oats. The farm of twenty-five acres which formed our previous illustration is Grade A, the most fertile of the three. The inferiority of Grade B soil as compared with Grade A is shown by the fact that the initial expenditure of capital and labor on B will return a smaller product than the same expenditure on A. Similarly the inferiority of Grade C to Grade B is shown by a still smaller product from the first expenditure of effort on this land. The relations of these different plots of ground to each other may be pictured in the following tables, the first of which it will be noted is repeated from our previous illustration.

GRADE A SOIL

Expenditure	*Total product*	*Marginal product*	*Average cost*	*Marginal cost*
$100	500	500	$.20	$.20
200	1,000	500	.20	.20
300	1,400	400	.214	.25
400	1,700	300	.235	.33⅓
500	1,900	200	.263	.50

GRADE B SOIL

Expenditure	*Total product*	*Marginal product*	*Average cost*	*Marginal cost*
$100	400	400	$.25	$ .25
200	700	300	.285	.33⅓
300	900	200	.33⅓	.50
400	1,000	100	.40	1.00

GRADE C SOIL

Expenditure	*Total product*	*Marginal product*	*Average cost*	*Marginal cost*
$100	300	300	$.33⅓	$ .33⅓
200	500	200	.40	.50
300	600	100	.50	1.00

It will be seen that the marginal cost for the first units of output is lowest on Grade A soil, next lowest on Grade B, and highest on Grade C, and that the principle of diminishing returns applies to the cultivation of each grade of soil, though the point at which it begins and the rapidity with which marginal costs rise as the

product is increased are not the same for all. In our analysis of these tables, one principle of universal application must be borne in mind; namely, that in no case will any produce be raised unless the selling price is high enough to cover the marginal cost of production. Regardless of how the ownership of the land is distributed — whether divided among a number of competing cultivators or concentrated in the hands of one — this principle will control its cultivation, determining the amount of produce obtained from each grade and from all of them together.

The appearance of economic rent. Now it is apparent at first glance that, until cultivation has been pressed beyond the point at which diminishing returns appear on Grade A soil, the other two grades of land will not be cultivated at all, since it will not pay to cultivate them. Until the third expenditure of labor and capital has been made upon the land of highest quality and the total output of that land has been advanced 1,400 bushels, the price of the product will not rise above twenty cents a bushel, this being the marginal cost. But the first units of produce on land of the second grade will cost twenty-five cents per bushel and on land of C grade, thirty-three and a third cents; to produce them therefore would involve a loss to the farmer which he will not bear. Up to this point (the third expenditure on Grade A) there will be no economic rent, even on lands of the highest quality, since the average and marginal costs of production are the same. If the price rises to twenty-five cents per bushel, it will then be profitable to make the third expenditure on Grade A soil, thus raising its output to 1,400 bushels, and at the same time to make the first expenditure on Grade B, the total product from this second grade being 400 bushels. Grade C will remain out of use, for the market price is still not high enough to cover its initial marginal cost. At this point there will be economic rent on Grade A, where cultivation has passed the point of diminishing returns, but none on Grade B, where marginal and average costs are still equal. Suppose that the market price rises again, this time to thirty-three and a third cents per bushel. A fourth expenditure will now be made on Grade A soil, a second on Grade B, while

the first cultivation will be made on Grade C. Economic rent on Grade A will rise; on Grade B economic rent will appear for the first time; while Grade C land, though under cultivation, will return no producer's surplus.

Why production is pushed to the limit. The question may be raised: How can we assume that cultivation will automatically expand with each rise in price until the marginal units of the product merely replace their cost? Why does not the farmer on Grade A or Grade B soil restrict his output so as to be able to sell all the units at a price in excess of their costs? The answer is that he will either not have the power to do this, or, if he has the power, it will be unprofitable for him to do it. If he is not the owner of the land, competition of other would-be tenants will cause a revision of the rent to keep pace with each increase in potential producers' surplus. The present tenant can retain possession of the land only by agreeing to bear these increased rent charges, after which, if he is to avoid loss, he will have no other alternative than to expand production to the point assumed. On the other hand, if he is the landowner it will be to his advantage to gain the maximum producers' surplus possible under prevailing market conditions; and this he can do only by advancing his marginal costs to equality with the price. Hence, as the thickening of population causes an ever increasing demand for farm produce and a constantly rising price, the history of agriculture in the little society we have imagined will be marked by two lines of development. There will be a growing intensity of cultivation on the better grades of soil and a wider spread of cultivation to take in soils of poorer qualities. Along with these changes will go rising rents on all lands which were formerly rent bearing and a widening of the rent bearing area to include soils which formerly returned no producers' surplus to their cultivators.

Intensive and extensive margins of cultivation. This illustration has prepared us to grasp the meaning of a concept of importance in the doctrine of economic rent: the concept of the *margin of cultivation.* The margin of cultivation may be defined as *that stage of tillage at which a unit of expenditure just pays for itself.*

This last, or marginal, expenditure is applied under such conditions that it produces an income only equal to the cost. That this is true may be due to either of two causes: the expenditure may be the last of a series applied to a given plot of ground, its product being restricted by the law of diminishing returns; or it may be the first expenditure made on lands so lacking in fertility or so remote from the market that they are barely worth cultivating at all. It is useful to distinguish these two cases by calling the one the *intensive,* and the other the *extensive,* margin of cultivation. The intensive margin of cultivation on any grade of land is reached when the product of the final unit of expenditure sells for no more than it costs to produce. The extensive margin appears on that grade of soil which it is just profitable to cultivate at a given time. On neither margin therefore is there economic rent. As we have seen from the preceding illustration, the position of both margins at any given time depends upon the selling price of the product. If anything occurs to raise this price, the intensive margin on all lands under cultivation will be raised and the extensive margin will be pushed out to lands whose lack of productivity made them extra-marginal at the old price. A fall in the price will, on the other hand, cause a lowering of the intensive margin on all lands and a narrowing of the extensive margin by the process of throwing the poorest lands out of use. It should be noted that these statements hold good only on the assumption that costs of production remain unchanged as the selling price fluctuates.

Practical illustrations of the intensive margin. The actual operations of agriculture afford abundant illustration of these concepts. The intensive margin is implied in the exploitation of every farm regardless of its size, its location, or its inherent fertility. The lands of the world differ in this respect only with regard to the amount of expenditure which can be made upon them before the no-rent stage of tillage is reached. Nearness to the selling point means a low scale of marketing cost which makes possible a relatively large expansion of product per acre. This is a determining factor in the case of perishable fruit and vegetable crops

and in the case of all forms of produce whose transportation costs comprise a large element in the selling price. Lands so situated as to be available for the production of such crops can be exploited to a high degree of intensity before the mounting marginal costs rise to equality with market prices. Similar in effect is nearness to the shipping points on railroads and waterways. The influence of this factor may be seen in the concentration of labor and capital on the truck gardens in suburban areas, as well as in the relatively more thorough cultivation of farms bordering on railroads as compared with those in back-country districts. Marked differences in the native fertility have very similar effects on relative intensity of cultivation. In many cases also a difference in intensity of cultivation, as measured by the number of workers per acre of land, is due chiefly to differences in the wages cost of the farmer. When men are cheap, more can be used before marginal expenditure equals market price. This factor also plays a part in causing the relatively high intensity of cultivation of the farms of Europe and Asia as compared with those of the United States.

The point must be stressed that the same principle governs the cultivation of all these lands. The fact that the application of this principle results in some cases in a large expenditure of capital and labor per acre and in other cases in a small expenditure is to be attributed solely to differences in the land — in its inherent qualities or its location. This fact is imperfectly understood by many people who write on the subject of agriculture. When the superior productivity per acre of the farms of Europe is held up as a reproach to the American farmer or the enormous produce of the small holdings of truck gardeners is made the basis of an argument in defense of the "small farm well tilled," the writers are exposing their ignorance of the fundamental importance of the intensive margin of cultivation. The truth is that, under present conditions of cost and selling price, more thorough cultivation of American farms is not profitable.

The extensive margin in practice. The concept of the extensive margin has similarly practical implications. Much of the total land area of the world lies completely outside the margin of culti-

vation for all possible uses. In the United States in 1920 no less than 40,000,000 acres of land were extra-marginal. Lands on the extensive margin are somewhat more difficult to identify since they are at the time under cultivation. According to definition, the product of the marginal lands returns an income just sufficient to cover costs of production. A slight rise in cost or a decline in sale price may be expected to throw these lands out of use. By applying this test, it is not difficult to convince ourselves that a vast quantity of land of this character is under cultivation at any given time. In a typical farming state like Iowa it is to be observed that almost all the larger farms contain plots of thin or intractable soils which are marginal land. These plots are periodically drawn into cultivation or thrown into idleness in response to variation in price and cost. In certain sections of the West entire farmsteads are to be found whose cultivators are making so little in the way of surplus that the slightest misfortune causes them to give up the struggle and abandon their lands. Sometimes these marginal lands are kept under cultivation only through the ability of the farmer to escape some ordinary cost of operation. In the New England States and in the Lake region many struggling farmers continue operations only as long as they succeed in avoiding payment of taxes. When the tax collector eventually insists on payment their lands are abandoned and, if later sold by the State, are apt to pass through a similar cycle under another cultivator.

The extensive margin is not the same for all the different uses to which farm lands can be put. Many crops demand certain qualities in the soil which are possessed by relatively small sections of the land area. For example, there are soils in the Connecticut River valley which produce a type of tobacco leaf very much in demand as wrappers for cigars. These soils are limited in amount and do not possess the requisite qualities in like degree. It is possible, under given conditions of price and cost, to define the lands which are marginal in this particular use; but lands lying outside this margin may profitably be turned to other uses. So it is with other crops which require peculiar qualities of soil and

climate. In other cases it is transportation cost which determines the marginal lands for certain uses; here too the lands which are extra-marginal from the standpoint of these particular uses may be well within the margin for other crops. Because of these complexities it is difficult to estimate the total quantity of land on the margin of cultivation in a country so large in extent and of so diverse climate and soil as the United States. We may however assume that pasturage is the lowest stage of cultivation — that is to say, the stage which involves the smallest expenditure per acre — and that lands unfit for this purpose lie outside the margin of cultivation for all uses. In 1920 there were 285,000,000 acres in the farm area of the United States which had never been devoted to any other use than this; and of this quantity, 135,000,000 acres were lying idle because of their inability to return the moderate costs involved in pasturage.

Economic rent equivalent to the contribution of land to the social income. It will be recalled that we set out upon our analysis of economic rent with this general proposition before us: in a competitive society, recognizing the institution of private property in land, each landowner will draw a money income equivalent to the contribution which his particular plot of ground makes to the common income fund of society. We have now explained how this income is brought into existence by the play of natural forces, how it rises and falls as the margins of cultivation change, and why it accrues to the landowner in the implicit form of producer's surplus or explicitly as contractual economic rent. In the course of this explanation, the equivalence of economic rent to the contribution of the land in question to the common fund of society was also established, though perhaps the point was not emphasized sufficiently.

When the land area of a country is parcelled out among several owners, it lies in the power of any one of them to reduce the social income by withdrawing his land from use. But the extent to which society suffers loss from this action is not measured by the total former product of the land, since holding this particular plot idle will set free labor and capital which will be productive

of income in other uses. This labor and capital however must now be applied on the margin of cultivation — either on the poorest lands at the time under cultivation or as the least effective units of expenditure on better grade soils — and consequently will be less productive than it was before freed by the action of the landowner in question. *The difference between the total product of a given piece of land and the product resulting from an equal expenditure on the margin of cultivation is the measure of the specific contribution of that land to the social income.* This is what society would lose if the land were withdrawn from cultivation by its owner. But it is also the measure of the economic rent of the land in question. We measured economic rent in the preceding discussion by the difference between the marginal and average costs of the producer — the marginal cost being equal to the selling price and hence determining his total income; the average cost determining his total expenditure. Marginal cost however is governed by the productiveness of labor and capital on the margin of cultivation; while the average cost, on a given piece of land, is controlled by the productiveness of the labor and capital as a whole on that land. Here then we have two equivalent expressions for the measure of economic rent; one viewing the matter from the social point of view and stating the measure of rent in terms of product, the other, from the point of view of the private individual, measuring economic rent in money terms.

Relation of economic rent to private property in land. In presenting this brief explanation of the relation of the two margins of cultivation to each other and to the price of the product, we have drawn an extremely simplified picture of the history of economic rent as a social income, from the time when lands first became subject to the institution of private ownership and the increasing pressure of population on land pressed cultivation beyond the point of diminishing returns. There has been a reciprocal relationship between economic rent and the institution of private property in land. In our view of the matter thus far it has been made to appear that the existence of economic rent is maintained in society because of the legal sanction which at-

taches to private ownership of land and other natural resources. This is true when we view rent as a form of private income. If land were not held by the individual members of society as a private possession, to be turned to the personal economic advantage of the owner, contractual economic rent would disappear; for no one would be permitted to dispose by bargain of a thing which was not legally his. Producer's surplus, or implicit economic rent, however is not created by a legal institution but by a law of nature — the law of diminishing returns. It must always exist so long as nature continues to reward with increasing niggardliness our efforts to wrest larger streams of supply from the resources which she affords mankind. Private ownership directs this implicit income into the hands of those who, by foresight or good fortune, have become possessed of parts of our natural resources which are more productive than those which lie at the margin of utilization. The institution of communal ownership might supply the foundation for a different disposal of this surplus gain — a roughly equal division of it among all members of society or some other arbitrary distribution according to moral or ethical precept — but it would not wipe economic rent out of existence.

However the causal relationship between private property and economic rent as a private income may, with equal correctness, be viewed as the reverse of that just presented. In all societies, even in one so populous as our own, there has always been an abundance of free land — land which no one considers worth even that trifling exertion which nowadays is involved in the process of preëmption and the registering of legal title. Turning back to our illustration of the different grades of soil, it is clear that the poorest of the three grades will be appropriated by no one as long as cultivation on Grade B has not reached the point of diminishing returns; Grade B, likewise, was free land until costs of production began to rise on Grade A. Similarly, if we assume such an abundance of the highest grade soil that labor need not be concentrated on any portion of it to the point which brings diminishing returns into operation, the choicest lands will also be a free

good; in other words, there will be no economic rent and no private ownership of land under these conditions.

This was, in general, true of our own country during the early colonial days. In all parts of the country, as it was known at that time, ample supplies of land of the quality best adapted to the needs of the time were available to all comers. No one would pay to another a rental for the use of a resource of nature which he could have for the taking; nor could any cultivator draw an implicit rent in the absence of the phenomenon of rising costs. There was, it is true, private property of a sort; for most lands could be made fit for cultivation only through an expenditure of labor in clearing forests, fencing, ditching, etc., and this resulted in an admixture of valuable capital with the land unit. But unimproved lands, even of the highest grade, were not the subject of private ownership; and this because they bore no economic rent. Broadening our point of view to take account of the sweep of the development of human society upon the earth, it appears that one of the most potent forces breaking down communal ownership under the tribal economy was the emergence of economic rent or producer's surplus upon certain portions of the tribal domain — a consequence of the appearance of diminishing returns under an increasingly intense utilization of the soil. Economic rent called into play the motive of self-interest which led the stronger or more fortunate members of the community to appropriate the better plots of ground to their own exclusive use. Legal sanction arose after the fact to endorse and support an institution which had begun to exist. And so the institution was maintained throughout the course of human history, broadening and perpetuating private income in the form of economic rent as the rapidly mounting numbers of the human race pressed with increasing intensity upon the resources of their natural environment.

Economic rent a regulator of agricultural production. Now that we understand the essentials of the law of economic rent, certain features of agriculture as an economic enterprise may be discussed. It is possible to place each variety of agricultural

produce in a class by itself and to work out the laws of rent for these different varieties separately. As stated above, the marginal lands are not the same for all these different uses. The production of some crops necessarily involves higher costs than others; some require lands of higher native fertility or of more efficient location with respect to the market. All the varieties of farm produce may be said to compete with each other for the use of any given piece of ground. It rests with the owner to decide, within the limits imposed by the nature of his soil and the climate, what kinds of crops he will produce, and his decision will turn on considerations of the amount of economic rent to be derived from the production of this kind of produce or of that.

Hence, for society as a whole, the controlling factor determining what kinds of goods will be produced, and what quantities of each kind, is the relative amounts of economic rent returned by the different utilizations of the land. If we imagine for a moment that farmers are wise, far-sighted, calculating business men, we shall expect each so to employ his holdings of land as to expand his producer's surplus to a maximum. He will first produce that crop for which the difference between marginal and average costs is the greatest, and he will continue this line of production until it ceases to return a surplus profit larger than does some other crop; at this point the production of the second crop will be begun, to be followed by a third and a fourth in the order of their profitableness. In this way the land is allocated to the different lines of cultivation so as to produce the highest profit possible under existing prices for the various kinds of produce. These statements do not imply that any single crop is grown continuously on a given plot of ground. The farmer, taking a long-run view of his problem, will sacrifice profit in a given year in order to increase his average or normal rent by preserving the fertility of his soil. Since rotation of crops is a basic principle in modern agriculture, the principle we have been discussing should be stated in terms of that *combination of crops* which returns the largest rent.

To this desire to obtain the maximum producer's surplus is also to be attributed the tendency of the market prices of many crops

to move in harmony with each other. If the price of one rises out of proportion to that of the others, the prospect of increasing his economic rent by devoting a larger area to the production of that particular crop will cause the farmer to abandon other types of cultivation. The quantity of the profitable crop will increase with a resulting tendency to check its rising price; while the quantities of the other kinds of produce will diminish and their prices will tend to rise.

At times in the world's history the changes in the relative profitableness of the different crops have been of sufficient magnitude to work something like a revolution in the agricultural organization of society. Long-continued production of certain crops and the practice of the form of cultivation peculiar to them produce in time a type of civilization which reflects the peculiarities of this underlying economic foundation. The introduction of a new form of cultivation will then be opposed by all the forces which rally to the defense of an established societal order — the pressure of public opinion, the weight of law, even the force of arms. Nevertheless, if the new forms are effective in increasing the economic rent of land, they will eventually displace the old. Such a change came on our own frontier within the memory of men now living when the immense ranches of the cattle " barons " were displaced after a bitter and often violent struggle by the fenced fields of the sheep raisers and of those who practiced mixed farming. England passed through a similar agrarian revolution at two different periods in her history, once in the sixteenth and again in the eighteenth century. These striking incidents in the history of agriculture serve to call our attention to the influence of economic rent as a force apportioning the arable soil to the production of various crops — a force whose effectiveness is none the less real in normal times because its operation passes unnoticed.

Some social consequences of diminishing returns. Throughout our discussion of the fluctuations of the margins of cultivation and the rise and fall of economic rent, we have assumed that the immediate cause was a change in the selling price of farm products. As explained, the law of diminishing returns can best be

understood on the assumption of a fixed point in time, and this implies that other variable factors — quantity and quality of land and stage of the arts of agriculture — remain constant. This is not to say, however, that a change in these factors will wipe out of existence the principle of diminishing returns; this principle will operate as before, but after the change it will operate from a new starting point and may result in costs which rise more or less rapidly than before. If these factors which control the costs of agricultural industry do remain unchanged over long periods of history, the natural growth of population will cause economic rent continuously to rise, create a larger aggregate of income flowing into the hands of the landowners, and raise to a position of peculiar economic privilege that class within society which discharges no function other than mere ownership. The increase of numbers will mean not only more mouths to be fed from the produce of the soil, but an increasing pressure upon the supporting power of the natural environment exerted through as many different channels as there are different branches of industry in the social order; for all of our industries run back in the final analysis to the land from which they draw their supplies of raw material. Because of the operation of diminishing returns, each addition to the population must set out to win subsistence under conditions less favorable than those which confronted their predecessors, even if the fertility of the land be maintained and the supplies of other natural resources prove inexhaustible. As the society grows older, its economic history will be marked by a tendency of the *average* income to fall and of the *inequality* of its distribution to grow more pronounced. The forces of diminishing returns, which work toward the general impoverishment of the group, will at the same time make for the relative prosperity of that class of society which holds the land in private ownership.

Students of history will recognize in this a key to the economic history of certain of the older societies of the world which have concentrated an increasing population upon a limited land area under stereotyped processes of production. It must be the fate of all expanding communities of men — our own no less than

others — which do not offset the effects of diminishing returns either by improvements in the arts of industry or by additions to their land area. To a large extent in human history escape from the impoverishing effects of diminishing returns has been found through the latter of these alternatives. The great migrations of races which marked the opening of the Christian era may be explained in this light.[1] Throughout recorded history one of the most potent causes of war between nations has been a difference in the birth rates of two groups of people on either side of a political frontier. The nation with a redundant population, feeling the cramping effects of diminishing returns in its basic extractive industries, is prone to look with covetous eyes upon the relatively more prosperous condition of its neighbor. Today this inveterate "land hunger" of nations takes the form of imperialism — a struggle for a "place in the sun," the place being found at the expense of the weaker and more backward of human races, whose land and other natural resources are appropriated for the relief of the peoples of higher culture. The greatest event of this kind in modern history was, however, the discovery of the New World with its great wealth of natural resources, which relieved the pressure of population on land in the Old World by opening up fields for colonization. Today no further relief of much consequence to human society can be expected from further increases in available land area, since almost the entire surface of the world has been appropriated by the dominant races of mankind. To offset the effects of diminishing returns, reliance must now be placed in the power of scientific progress to advance the arts of industry to higher stages of efficiency.

Every improvement in the technique of production, like every addition of fertile lands to the accessible supply, not only has this broad social effect of raising the average prosperity, but, what is more directly related to our present problem, it results in a lowering of economic rents. The immediate effect of such a change is to introduce new conditions of cost into agricultural industry. A

[1] Impoverishment of the soil through wasteful cultivation was also one of the causes of these migrations. This is not the same thing as diminishing returns.

given expenditure of labor and capital under the improved conditions will result in a larger product, and this in turn will tend to depress the market price. Using more efficient methods, society can produce a larger aggregate output of goods before it feels the effects of diminishing returns on the most fertile lands. If we assume that the demand for the product of our farms remains unchanged, the marginal lands will be thrown out of use and the intensive margin of cultivation will be lowered, with a consequent fall of economic rent to landowners. If, on the other hand, we assume that the demand increases while this change is occurring — and this assumption more nearly conforms to the facts in our expanding society — we shall find two counteracting influences at work upon economic rent. The tendency of an increasing demand for the products of agriculture is to raise economic rents and enlarge the segment of the social income which goes to the landowner. A reverse tendency results from the improvement in the technique of production; and these two tendencies serve to offset each other.

Economic rent not a cost of production. If our discussion of the laws of rent has been fully understood, it has become clear that economic rent is not one of the costs which determine the selling price of agricultural products. This is a matter concerning which there is much confusion of thought. The statement seems to conflict with the experience of every tenant and is, therefore, frequently condemned by the practical man as contrary to the facts of everyday life. To the tenant, the payment of an annual rental constitutes one of his costs of production, a cost no different in nature from the payments he must make for capital and labor. He includes this cost in his reckoning when calculating the profitableness of any form of cultivation and expects to recover it from the sale of his produce. From his point of view, economic rent is one of the factors which govern his calculation of profitable selling prices.

Now it is true that there is a relation between economic rent and price, but the causal connection runs from price to economic rent and not from economic rent to price. It is because the price is high enough to return a large producer's surplus that a high rental

is exacted for the use of a given piece of land; it is not the payment of the high rent which causes the price to be high. The truth of this statement may be substantiated by a simple analysis of the relation of the supply of agricultural produce to its price.

The cost of the *marginal units* determines the selling price of a given quantity of produce. We can discover this marginal cost on both the intensive and the extensive margins of cultivation; but it will simplify the problem if we first focus attention on the extensive margin, since this, consisting as it does of a certain type of land in its entirety, presents a less difficult subject for exposition. The cost of production on this marginal land then controls the selling price of any given kind of produce. But there is no rent returned to the cultivator of the marginal land; hence economic rent cannot be one of the elements making up that cost which determines price. This land was drawn under cultivation because the price of the product was high enough to justify the expenditure of capital and labor upon it. By working it, the owner receives from the sale of his product an income sufficient merely to cover his cost of production — wages, interest, and the normal return for his own effort. To demand an arbitrary rental for the use of this marginal land and to attempt to recover this additional cost by raising the price above what the buyers are willing to pay for the entire quantity of product brought to the market would be futile. The higher price would make it profitable for the cultivators of rent-bearing lands to increase their output, and this would still further increase the quantity supplied in excess of the amount which buyers were willing to take at the price set. The lack of a market for this excess quantity would soon depress the price to the point where it no longer exceeded the costs of production on no-rent lands.

It is sometimes said in opposition to the statement that economic rent is not a price-determining factor of production that there is, in actual practice, no land under cultivation which does not return a rent. This assertion is contrary to the facts; but it may be granted without affecting the point at issue. Let us assume that, in our country, all lands are under cultivation and that every plot

returns an economic rent. How could such a condition come about? By reason of one thing only; namely, that on all soils cultivation has been carried beyond the point of diminishing returns, with the result that every cultivator sells his produce at a price which stands above the average cost of production. In such a case, we could discover the marginal cost of production which determines market price only by referring to the intensive margin — that unit of expenditure which, on each grade of land, it is barely profitable to make. But here we would arrive at the same conclusion regarding the causal relation between rent and price; for the product of this expenditure, which is the price-determining unit of the supply, returns no economic rent. On the intensive margin, average and marginal costs are equal; no producer's surplus accrues to the landowner as the result of this expenditure.

Rent and cost of production in practice. The general rule that rent is not a price-determining cost is one of the laws of economics which "practical" men find difficulty in accepting. This is because it appears to conflict with business experience from two points of view. In the first place, when we view the problem from the standpoint of any particular crop, it appears that the potential rent of a given plot of ground in another use may properly be considered one of the costs of producing the crop in question, provided this crop returns a lower rent than would have been returned by the alternative use. To illustrate, suppose a sixty acre field will return $600 when planted to corn, $300 when planted to hay, and $120 if used for grazing. Under these conditions one of the costs of raising hay on this land will be the $5 loss of rent per acre involved in the choice of this less productive type of cultivation; and the rent cost for pasturage will be $8 per acre. But these costs of the alternative use can only in rare circumstances enter into the price of hay or meat products. Ordinarily they will be prohibitive costs as far as the less productive crops are concerned. Corn will win out in the competition for the use of the land, and the other crops will be relegated to soils on which they are not at a disadvantage as compared with corn. If any individual farmer

sets aside this land for the production of hay or meat, he may rightly consider the alternative rent a cost of production and add this to other costs in computing a profitable sale price for his product. But there is no reason to believe that the price determined in the market will rise to cover this cost.

In the second place, rent appears to be a cost of production when considered from the point of view of the individual farmer. At other points in our study, we have seen that causes and effects often stand in inverse order according as they are regarded from the point of view of the individual or from that of society as a whole. For example, the price of coal at a given time is the effect, in part, of the quantity of coal bought at that price by a given individual. But to this individual the price is one of the causes of his having bought that particular quantity of coal. So it is with contractual economic rent. The rents which tenants must pay for different plots of ground are determined, as explained, by the amounts of producers' surplus which the prices of their products make possible. When a tenant has obtained the use of one of these plots by contracting to pay this rent, he very naturally considers the rent a cost which must be recovered from the sale of his product if his business is to prove profitable. But if he is unable to obtain this profitable price and, therefore, unable to recover the rent cost, he will be compelled to surrender the land to some other farmer whose energy and skill are superior to his own. And if no other can be found who can sell the product of the land at the existing price and recover the rent in addition to his other costs, the land will be kept in use by lowering the rental, not by raising the price of its product.

Viewing contract rent as a cost from the standpoint of the individual farmer brings to light certain inferences of rent theory. When lands are cultivated under lease, the obligation to pay rent to the owners has the effect of equalizing costs of operations for all the tenants, in so far as these costs differ because of differences in the quality or location of the lands. Since the rentals absorb all the producers' surplus, the tenants on the more fertile lands or on those nearest the market have no advantage over their less

fortunate competitors. This is not to say that all tenant-farmers make the same profit; or that those on the fertile lands of Iowa are not more prosperous than those struggling with the stony hill sides of New England or the intractible soils bordering on the desert of Nevada. Such statements would be absurd in the face of common observation that the reverse is true. Farming is a business enterprise like any other form of production for the market; and farmers differ in their skill and efficiency as managers. The rent contract does nothing to equalize these differences in personal ability of the tenants; it merely sets up conditions which all tenants must meet in order to win access to the land. The abler farmers would make more money than the less able, even if both groups were placed side by side on lands of the same quality. But in practice the abler farmers tend to get possession of the better grades of land — those which command the higher rentals. This is because these lands afford the widest opportunity for the display of the productive ability of the tenant; while their high rentals act as prohibitive costs to the ignorant or wasteful farmers. Hence we find, as a rule, that the relative prosperity of the tenants varies directly with the quality of the land they occupy; but this is due to differences in their own capabilities. Stating this aspect of rent theory with greater precision, we may say that contract rent tends to equalize costs of production for tenants of equal ability.

CHAPTER XXXV

RENT OF URBAN SITES AND MINES. PROBLEMS OF RENT

Producer's surplus or economic rent on urban lands. The conclusions reached in the preceding chapter hold true of natural resources in general, including the ground area embraced within the limits of our cities and towns. The problem of economic rent is complicated in the case of urban lands by the diversity of uses to which these lands can be put and the intangible nature of their products. To simplify the problem, let us consider first but one of the different uses — that of the retail merchant. It is obvious that the several portions of the land area of the city differ in their effectiveness to serve the purposes of the merchant. It is possible to throw the available sites into a series of classes ranged in the order of increasing effectiveness from the lowest to the highest.

Let us suppose now that merchants of equal ability and with equal amounts of capital at their command have built retail shops upon all these sites, and that these merchants are in competition with each other. Obviously the competitors will prosper in varying degrees. The annual income of the merchant with the best location will exceed that of his neighbor on the second best site; this second merchant in turn will receive a larger net income than his competitor with the next best location; the income of the third merchant will exceed that of the fourth; while the merchant whose shop is situated on the marginal land will struggle along with an income barely sufficient to keep his business alive. Nor can these differences of income be attributed to differences in the efficiency of labor and capital employed, for we have assumed that these factors are equal in quantity and quality in all cases. They must be due solely to the varying efficiency of location of the different

shops; and this is an attribute of the land on which the shops stand. The income of the merchant on the marginal site, small as it is when compared with the incomes of his more fortunate competitors, is nevertheless sufficient to cover all his necessary costs of production, including his own normal reward as an enterpriser; otherwise this man would withdraw from his present business and strike out on some different and more promising venture. But if this is true, then the excess of income enjoyed by the other merchants must be viewed as producer's surplus; in other words, all the competitors except the one who owns the marginal site receive implicit economic rent.

Here we have indicated, in a rough way, the nature of economic rent on urban lands, on the assumption that the lands are devoted to one particular sort of business undertaking. Similar differences exist with regard to the profitableness of different sites from the point of view of any other type of business that we might choose to mention. The theatres and hotels, the office buildings and apartment houses, the banks and wholesale establishments in any city of considerable size occupy sites which, even to the most casual observer, differ in respect of their adaptability to the purposes of the business. In most cities, this difference is so noticeable that it is not difficult to point out the enterpriser who has surpassed his competitors in the scramble for desirable locations. It is often a matter of common remark that the owners of the Wonderland Picture Palace, for example, are so situated as to get the cream of the profits in the moving picture business; or that Jacobs and Goldstein have the choicest location in the tailoring world; or coming down to actualities that United Cigar Stores are to be found on the busiest corners. We need no profound knowledge of business to conclude that each of these leaders in his own field would sacrifice income if he were forced to move to a less desirable location, even if he applied his labor and capital as abundantly and efficiently as before. There is no way to explain these differences in the income-producing powers of enterprises differently situated except to attribute them to the relative productivity of the lands on which they stand.

Diminishing returns from urban lands. Sites of varying productivity are used simultaneously for a given purpose because of the operation of the law of diminishing returns. Diminishing returns on urban lands, for many of the lines of use to which they may be devoted, do not appear until utilization of the land has reached a point of high intensity; but they do appear sooner or later. An increased expenditure of capital and labor on a site of given dimensions results in the addition of more stories to the building, a larger area of floor space, an increased stock in trade, increased working force, a more complex and unwieldy organization. The gigantic department store of the modern city shows to what extent this increasing intensity of utilization may be pushed on the best sites before rising costs of operation wipe out the advantage which the site possesses as compared with others less well situated. But this expansion cannot go on indefinitely without eventually causing a diminution in the income derived per unit of labor and capital expended. As the building grows in size, the costs of maintenance, of transportation of goods and passengers from floor to floor, of supervision, will eventually rise. Many small sources of waste will pass unnoticed by the managers which in a smaller establishment would be discovered and corrected. For many reasons the costs of rendering the service will rise, or, to use the correlative expression, the returns from the last unit of expenditure will fall. Hence the time must come when additional expenditure on the best sites no longer affords a larger profit than the same expenditure on lands not so well situated.

Diminishing returns and the margins of utilization. At this point the enterprise will begin to spread out. If the owners of the best sites are also owners of the next best, it will be to their advantage, instead of investing more capital in their present establishment, to begin the development of the other land. But if this is not done by the merchants already in business, competing enterprises will be set up by others; for now the advancing costs on the best sites will have made additional expenditure here no longer more productive than is the initial expenditure on less efficient lands. As the process of expansion goes on, diminishing returns on the

second site will ultimately reach the point where it is worth while to begin utilizing the land which ranks third in order of productiveness. A similar tendency on the third site will draw a fourth into use, and so on.

Here we have illustrated the concepts of the intensive and extensive margins of utilization on urban lands used for retail merchandising. The intensive margin is that stage of exploitation — typified by an added story, department, operating device — which it is worth while to operate at a given time on any particular site of given size; the extensive margin appears on the least productive site, which it is worth while to use for retail shopkeeping. No broadening of the extensive margin to include less productive sites will occur as long as marginal costs on the lands already in use are lower than the initial costs on the extra-marginal lands. The returns on the two margins tend to keep pace with each other. Each broadening of the extensive margin will be accompanied by an increased expenditure in the form of higher buildings and larger equipment on the better sites — that is to say, by a raising of the intensive margin.

When economic rent from urban lands is an implicit income — that is, when the owner of the land is also its user — the problem of estimating its amount on a given site is extremely difficult. By comparing one site with another, as we have done in the preceding discussion, we can assure ourselves that rent exists; but in any given case it is so inextricably merged in the total net income which the owner derives from the operation of his business that he himself is not able to mark it off to the last dollar. But in actual practice its appraisal is not left to the owner to be worked out by an analysis of his money receipts. In every city of average industrial activity the amounts of economic rent on various grades of land are measured with a degree of accuracy sufficient for the purposes of the business world through the activities of speculators in real estate and the competition of tenants for the use of land they do not own. These activities result in a schedule of explicit rents, or contract payments for the use of land, which very closely approximate the amounts of producer's surplus which

the different plots of ground would return if exploited by their owners.

Urban site rents in other uses. We have simplified the problem of economic rent from urban lands to an artificial degree by limiting attention to one of the many possible uses to which the land might be devoted. For each of these uses of urban land there is a set of conditions which tests the relative fitness of all the available sites to serve that use. A summary of these tests in certain uses is given by a leading authority on this subject as follows: "In the case of retail wholesalers and light manufacturing, proximity to the retail stores which are their customers; in the case of wholesaling or manufacturing, proximity to transportation; and in the case of public or semi-public buildings, for historical reasons, proximity to the old business centre. Hotels of various classes seek locations similar to the retail stores of the same classes on convenient traffic streets which advertise them. The highest class apartment hotels seek locations on or near such traffic streets as run through or near the fashionable districts. Below this grade the various classes of flats seek locations for the convenience of their tenants, tending to draw nearer and nearer their tenants' places of business until finally we reach the tenements crowded among the factories where their occupants work. The basis of residence values is social not economic, the rich selecting the locations which please them, those of moderate means living as near by as possible, and so on down the scale of wealth, the poorest workmen taking the leavings, either adjacent to such nuisances as factories, railroads, docks, etc., or far out of the city. Certain features appear to attract the wealthy in selecting their residence district, among these being nearness to parks, a good approach from the business centre, not too far nor yet too near, a moderate elevation if obtainable, favorable transportation facilities despite the fact that the rich ride in their own carriages and automobiles, and above all absence of nuisances." [1]

It would be possible by extended analysis to search out the attributes of land which affect its productiveness in each of the many

[1] Adapted from Rickard M. Hurd, *Principles of City Land Values*, 1903.

uses to which it may be put in the life of the city; but this is unnecessary for our purpose. Whatever the test applied to determine the adaptability of the land to these different uses, it is obvious that the sites will vary in effectiveness from the standpoint of any one of them. As in the case of retail trading, so in the hotel business or in banking or in manufacturing, some plots of ground will offer a maximum reward for a given expenditure of capital and labor. In each use another group of sites will rank a little lower in effectiveness than the best; others, lower yet; and so on down to that plot of ground — the marginal site — which it is just worth while to utilize for the purpose in question. Moreover the principle of diminishing returns will apply to each group of sites regardless of the use to which it is devoted. Hence, if we focus attention upon any one of the possible uses of city land, we shall find exemplified the same law of economic rent governed by a common principle of diminishing returns — the rent of any given plot of ground being determined in each case by the superior effectiveness of the land in question as compared with the marginal land.

Comparative economic rent a governing force in the location of urban enterprises. We may conceive of all the varied activities of the city as being in competition with each other for the possession of the city's land area. The distribution of this land among the different uses — in other words, the location of the city's enterprises — is determined by a law of *comparative economic rent.* The self-interest of owners and tenants works constantly in the direction of assigning each plot of ground to that use in which it will produce the maximum income. The theatre district, the shopping district, the hotel district, the residential district, and so on, are given their distinctive character in any city by reason of the fact that these areas of land happen, for the time being, to be more productive in their present uses than in any other. When land is offered for lease, it is the enterpriser for whose particular purpose the land is most productive who can afford to pay the highest rental for it. This applies with equal force to lands offered for sale; for, like agricultural lands, urban sites derive their value from the discounted economic rent, and the man

under whose management the rent will be largest can afford to offer the highest price for it.

In a growing city this functional distribution of land never remains stable for any length of time. A plot of ground which today, when subjected to the test of comparative economic rent, proves to be most productive as the site of a retail shop may presently come to have superior effectiveness in some other line of business. These changes in comparative rent are often the product of trends of development in the city's life whose influence is largely unforeseen. An alteration in the street car system, a new location of the transfer points, the opening of subways and suburban roads, the invention of a new device for rapid transportation, will alter the effectiveness of the different areas of the city as locations for the city's various enterprises. To select an example, the popularization of the telephone has exerted a profound and far-reaching influence on the functional distribution of urban sites. Such an occurrence as the opening of a new railway terminal or the reservation of land for city parks, may destroy the balance of comparative productiveness among the neighboring sites and even make its influence felt in distant areas. It is said that the construction of the Brooklyn Bridge by diverting traffic from the old Fulton Street Ferry and throwing it half a mile back from the river on either side, removed millions of dollars of value from the streets leading to the ferries, especially in Brooklyn. The history of any thriving American city will furnish abundant testimony to the influence of such changes on the functional distribution of land. What was once a neighborhood of private residences may in the course of time become a shopping district; the shops may in turn give place to wholesale establishments; these to factories or to gas works or to railroad freight yards. This shifting of the city's activities from one district to another is but another illustration of the outworking of the law of comparative economic rent whose tendency is to allocate each site to that use in which it is at the time most productive.

Apparent exceptions. Certain facts which may appear as exceptions to the laws of rent are to be explained as consequences

of this continual changing of the relative effectiveness of urban sites. In strict theory urban lands should be arranged in concentric circles around a number of nuclei or focal points, each representing the most effective sites in one particular line of use. Lands outside the margin would not be used for urban purposes; they might of course be devoted to agriculture. But all plots of ground within the margin would be utilized in degrees of intensiveness which varied directly with their rent-producing possibilities. In almost every city however are to be found plots of ground, well within the margin of utilization, which are nevertheless allowed to lie unused. In most cases this is due to the changing conditions of which we have been speaking. It is poor economy for any owner of land to commit it to a given use by the erection of a permanent building, even if this use happens at the time to be most productive of economic rent, when conditions in the city are changing so rapidly as to indicate that presently some other use may become more productive.

In theory also the largest and most costly improvements will appear on lands which return the highest rents and successively smaller expenditure on inferior sites down to the margin. A glance at the buildings in the different sections of any city will convince the observer that this is in general true. But often one finds a high rent site occupied by a cheap or obsolete structure, or conversely a costly building erected upon a low rent site. This too is usually the consequence of changing conditions. When comparative productivity is liable to sudden changes, the same caution which induces some owners to allow productive lands to lie idle will cause others to hesitate to erect anything but temporary buildings on their land, lest it be committed irrevocably to a use whose profitableness is short-lived; hence the occasional high rent site imperfectly developed. The opposite of this — the expensive building on a low rent site — when it is not merely a monument to the stupidity of some speculator, is either the aftermath of a change which has already occurred or a shrewd anticipation of impending change.

Economic rent not a cost governing price. In our study of economic rent from agricultural lands we learned that rent is not

one of the costs of production governing the price of the product. This principle applies to urban lands as well; but here, because it runs counter to prevailing opinion, the true relationship between rent and price is somewhat more difficult to grasp. The problem is complicated by the fact that a given commodity sold in shops occupying different grades of urban land does not command a uniform price as is the case in agriculture. The same pair of shoes may be bought for ten dollars at one shop, for nine dollars at another, for seven dollars and ninety-eight cents at a third. The cuisine of a medium-priced restaurant may be superior to that of a fashionable place where one pays a cover charge of one dollar and acquires in exchange the privilege of being rapidly relieved of his money by check-room, waiter, and cook. Jones and Fink, department store owners, may display feminine headgear which is, in all regards except price, exactly like that of Madame Marie. The tailor's name and the extra twenty dollars in cost may be the only attributes which distinguish the clothes made on Fifth Avenue from those of the less fashionable firm around the corner. The belief that these differences in prices are caused by differences in rent is fostered by many business men, and especially by the merchants on the low rent lands, who, for advertising purposes proclaim ability to undersell their competitors because of their lower rentals. Certain firms of tailors who have adopted the policy of occupying only the upper floors of buildings use as their advertising slogan "Walk upstairs and save ten dollars." Shoe dealers, provision merchants, grocers, and others whose shops are located on side streets or in the less reputable neighborhoods, often make the same appeal to the frugality of the buying public. The merchants in the fashionable districts also rely on this widespread notion that high rents cause high prices as a defense when the customer complains that he has been overcharged.

The fallacy in this reasoning may be traced to a misapprehension of the product of urban lands. The product of land devoted to retail trading does not consist in the commodities offered for sale in these shops, but of an intangible service. The reward of this service is the income (commonly called profit) derived by the mer-

chants whose function it is to perform it. Any attributes of a site which increase this reward for a given expenditure of capital and labor in the retail trade will bestow a high rent on the site. The amount of the rent, it should be noted, is a consequence of those peculiarities of the land which enlarge the profit of the merchant.

High rent and high prices. In some cases a given site will increase the merchant's profit for no other reason than that it places him in position to overcharge his patrons. There are people in the world to whom indifference to thrift is a matter of pride; they will pay high prices without a murmur if by so doing they acquire merit in fashionable society. In every city of considerable size there are districts to which people of this type resort in large numbers to do their shopping, districts whose selection is the outcome of a multitude of factors — tradition, habit, accessibility, pleasant environs, etc. Upper Fifth Avenue in New York, Bond Street in London, Rue de la Paix in Paris are famous for this character. Access to such lucrative patronage is limited to those merchants who gain control of the limited areas of land in the fashionable neighborhood, and competition among merchants for the possession of this land causes its rentals to rise. Here there is a relation between rent and the selling price of goods, but the causal connection runs from price to rent, not the reverse. The immediate cause of the high rent, as always, is the high income; the income is caused in turn by high prices.

High rent and low prices. But the selling price of his good is only one of the factors affecting the enterpriser's profit; another factor of equal importance is his cost of operation. A merchant's profit will expand if his costs fall while his selling price remains constant, quite as effectively as if his price had risen while his cost remained unchanged. Moreover, if his lower operating cost is due to the character of the site upon which his shop stands, and not to his own peculiar genius as a manager or to the efficiency of his capital equipment, it is the land which really produces his extra profit and the owner of the land will absorb the excess in the form of a high rent. Among mercantile enterprises, large volume of turnover in comparison with necessary overhead expense is the primary

cause of a relatively low cost of operation. Sites on crowded thoroughfares and at the transfer points of traction lines permit of a volume of business so much in excess of that done by shops in less effective locations that the merchants who possess them can make abnormal profits at prices no higher than those charged by their competitors. In such cases, high rents will be associated with low selling prices of goods; though it is still correct to say that the high rent is conjoined with a high reward for the merchant's service; *i.e.*, with a high income or profit.

There are many illustrations of the effect of this factor. For example the sites in New York at Broadway and Thirty-fourth Street and at Madison Avenue and Fifty-ninth Street, in New Haven at Church and Chapel Streets, in Denver at Fifteenth and Lawrence Streets have access to an abnormally large volume of traffic. According to a Cleveland table of relative land values, a corner lot is worth 75 per cent more than an inside lot on the main street. Another factor increasing the value of a site for retail purposes is ability to display goods during the hours when shopping is done. Lots on the side of the street which is shady between the hours of eleven and four o'clock are normally worth from 20 to 40 per cent, and occasionally 100 per cent, more than those on the sunny side. "In southern cities where shade is even more important, the relative value of the four corners of two intersecting streets is well defined; the southwest corner being most valuable, the southeast next, the northwest next, and finally the northeast corner."[1] These illustrations refer to the business of retail merchandising; but the student will have no difficulty in applying them, with necessary alterations in detail, to the relative profitableness of sites for other urban uses.

The point to be emphasized then is that high rent is caused by those attributes of the site which enable its possessor to derive a large income from its use. Ability to charge high prices is only one of these attributes, and a relatively unimportant one. The buyer in any city will find that low prices are as frequently to be found on high rent lands as are high prices. In fact, it is upon

[1] See R. M. Hurd, *Principles of City Land Values*, 1903.

these lands that the so-called "cut rate" shops flourish. The United Cigar Stores, for example, are located only at the corners of busy streets; the five- and ten-cent stores, the chain drug stores, the large department stores and other enterprises which advertise a low price policy are similarly located on expensive land. It is true of course that any individual merchant who has contracted to pay the large rental of a neighborhood where the only aid to large profit is ability to overcharge his patrons must maintain his prices at a high level or fail. To him the necessity of paying the rental appears as a compelling cause of high prices, and he will explain the matter in this light to his patrons. But this is merely an illustration of the reversal of cause and effect which we have discussed in connection with the problem of rent from agricultural land.

Rent of mines. The income from mines constitutes no exception to the laws of economic rent. Mining is an industry subject to the principle of rising costs. If we examine the operation of any typical mine — a coal mine, let us say — it will become clear that the product cannot be indefinitely extended at a uniform cost per ton. As continued attempts are made to expand the output of the mine, it will become necessary to work less productive veins, to follow the galleries further at an increased expense for pumping, hoisting, and other facilities, to increase the amount of labor in occupations not directly contributing to the production of coal; in other words, the marginal costs will begin to rise. Up to this point, for reasons now sufficiently understood, no economic rent will appear in the mining industry; the owners of mines can exact no royalties in return for the concession of exploiting them. In our own country, however, and in all others where the mining industry has grown to any considerable proportions, the growth of the industry has long since passed the point at which diminishing returns appear. Mines of varying degrees of productiveness are being exploited simultaneously. The price of the product is high enough to cover the cost of the most extensive mine, and all owners of mining properties of a productivity superior to that of the marginal mines are in receipt of economic rent. The mining

of copper, iron, the precious metals, and other minerals is governed by the same laws.

Apparent differences between mines and other natural resources. Attempts have been made to differentiate mining from the other extractive industries and to disprove the application of the laws of economic rent to this group of enterprises, because of certain practical attributes of the mining industry. Mining is notoriously a hazardous undertaking. It has been likened to a lottery in which prizes are intermixed with blanks and neither reward nor failure bears much relation to the effort expended. Especially in the production of precious metals is prospecting largely a matter of chance. The returns of copper mining are only slightly less hazardous. If these enterprises were pure gambles the relation between the price of the product and the marginal cost of production, which is fundamental to the existence of rent, would be so slight and uncertain that our laws of rent would not apply. Such has been the case in previous ages with regard to the mining of precious metal, and it still holds true of many speculative ventures in this field and in others. But in recent years the mining industry as a whole has been organized on as stable a foundation as any other great enterprise. Today the operators of the larger properties determine, by a process of nice calculation in which cost is compared with price, to what point it is profitable to exploit their mines under given market conditions. When organized as going concerns, the difference in productivity between different mines is clear enough to the business world; and it has become a recognized practice to pay the owners of the land a "royalty" — the technical term denoting economic rent in the mining industry — which measures with rough exactness the producer's surplus of the mine.

Another apparent difference between mining and other extractive industries lies in the fact that a mine represents a store of wealth which will necessarily be exhausted after years of use. The mine's product is not a reproducible good or service like that which results from the utilization of agricultural or urban lands. In terms of accounting, a mine is a "wasting" asset; the mine itself

is sold piecemeal to the consumers of the product. Hence, it is said, the difference between selling price and average cost of production does not represent a producer's surplus but is a payment for the capital value of the wealth which is destroyed through the operation of the mine.

Now, however important this concept may be from the standpoint of the accountant, it is of no importance in connection with a study of rent. The mineral deposits of the earth, when considered from the point of view of society as a whole, are gifts of nature. They differ from those parts of society's economic equipment for which man is responsible — such as ships, machines, houses, railroads — in that they are not and cannot be produced by man; hence their supply is in no wise affected by the prospect of recovering an original cost from the use to which they are put. Capital instruments made by man will not be supplied unless the price of their services is high enough to return this original or "fixed" cost; but coal mines would be operated if the price of coal were high enough only to cover the cost of operation, leaving nothing as a replacement of the mine itself, which is exhausted in the process. Moreover, the price of coal will not, under competitive conditions, rise above this point which covers operating costs until the demand for coal presses the mining industry beyond the point of diminishing returns. When that time comes, the price will be governed by the marginal cost, and a surplus, or economic rent, will appear.

Diminishing returns a principle of general applicability. Throughout this study of the income from land and other natural agents, we have discussed the law of diminishing returns as if it applied only to man's utilization of the resources of nature. But this is a partial view of the matter; for a similar principle may be observed in connection with the other productive factors, under appropriate conditions. The condition essential to the appearance of diminishing returns is that one of the factors be held constant while those which coöperate with it are increased in amount. In our discussion thus far, the constant factor has been land and the variables labor and capital. As the variable factors increase in

amount, the output of the combination will diminish when measured in proportion to given units of the increasing factors. This statement could be made with equal truth if any of the three factors is held constant while the other two increase. Diminishing returns, in other words, is a principle of universal application, always showing itself in connection with any phase of man's economic activity when certain factors which comprise the business unit outrun the others.

To illustrate, let us assume that the constant factor is not land but capital instruments, machines of a peculiar type which are absolutely fixed in quantity. These machines are set up in a factory as an indispensable equipment for producing a certain commodity. For some reason strenuous efforts are made to increase the output of this commodity by the only method possible in the circumstances; *viz.*, by applying increasing quantities of the other productive factors to the exploitation of these machines. We can construct a mental image of the process by which the volume of output is expanded; the number of workers who operate each machine increased from one to two, three, four, five; the floor space allotted to each also increased with a corresponding expenditure of land and capital upon the productive process; managerial ability increased by a multiplication of managing officials striving to promote efficiency and prevent waste. In response to this growing intensity of effort the peculiar product of the machines would certainly continue to increase, at least up to the point where the workers became so numerous as to hamper and obstruct each other. But would the product increase in proportion to the effort expended? Would the third, fourth, and fifth worker tending each machine add as large an increment to the output as did the first? Would the later additions of floor space, the second general manager, the additional foremen and supervisors, add proportionately to the product? Certainly not; the time will come when the increase of product will fail to keep pace with the increase of effort expended. This is diminishing returns from the utilization of capital instruments; it is analogous, as we can see, to the principle which governs the utilization of land.

To vary the illustration, let us suppose that labor is the constant factor. Assume that in a given community only a few men know how to make the gunpowder upon which the community relies for military purposes, hunting, and the performance of many industrial processes. An emergency occurs which makes it imperative that the supply of gunpowder produced per day be greatly increased. To make the picture realistic, we may suppose that the community is threatened with invasion by its warlike neighbors. In the face of this emergency, ways could be found of expanding the output of this vital product over the volume of output normal in times of peace; but these ways would consist in increasing the amounts of the productive factors which coöperate with the fixed labor supply. Each craftsman would be supplied with additional tools, larger areas of land upon which to expand the scale of his operations, larger supplies of raw materials, increased quantities of auxiliary labor. Under these facilitating circumstances, the daily product of each craftsman, or of all of them combined, would increase. But the increase could not continue indefinitely to be proportionate to the increasing units of the cooperating factors. Exploitation of a fixed labor supply will show diminishing returns at some point in the process of expansion.

Applying the test to a particular kind of labor, the managerial function obeys this same law. If we have to do with a fixed number of enterprisers — men who possess the secret of management — and each is equipped with increasing quantities of land, labor, and capital instruments, the output of the productive unit represented by this combination of factors will increase. But after the business unit under the command of each enterpriser has attained a certain size, further expansion will result in an addition to the product less than proportionate to the additions made by earlier units of capital, land, and labor of equal size. These illustrations are admittedly extreme cases, intended to throw into sharp relief the operation of the principle of diminishing returns as exhibited by productive factors other than land. The principle of diminishing returns, to repeat, applies to each of the productive factors when the one in question is fixed in amount while the others are

variable. We shall see in the course of our study of distribution how important is this principle in the determination of interest, wages, and profits.

Features which distinguish rent from other income. But there are sufficient reasons why the rent income should be distinguished from these other types and why its existence raises peculiar social problems. There is no productive factor, other than land, which cannot be increased in amount over long-run periods, either in its entirety, or in its special forms. The whole mass of the laboring population can be, and normally is, increased by an expansion of the birth rate, a fall of the death rate, or both, and as for the fractions of the whole labor supply which make up the different occupational groups, they expand and contract under the influence of training and education. There is no skilled group, however refined its art and secret its mysteries, which can long defend a position of peculiar advantage in respect of income from the competition of oncoming generations of workers potentially equipped in mind and body to duplicate the service of the favored group. Continuous redistribution of the labor supply tends to prevent one group from long enjoying a differential advantage as compared with other groups. There are, to be sure, formidable obstacles to be overcome in this process of redistributing the supply of labor, and some of these obstacles have proved insuperable up to the present time. But at any rate the tendency is there, and its results may be observed in the fact that the wage rates of different groups of workers bear certain more or less fixed ratios to each other.

This is even more true of capital instruments. Through the voluntary savings of its members, society's total equipment of these instruments is increasing as the years pass, and may be increased at any given moment by making special appeals to the savers or offering special rewards to investors. No special type of capital instruments can be held permanently fixed in supply. When the income enjoyed by the owner of a particular machine begins to outstrip that of owners of other types of capital instruments, equal in cost of production, new machines of this special type will be made and set up in competition with those now in existence.

Capital flows from one form to another, seeking out the points of maximum income. This flow may be delayed or retarded by special monopolistic devices, such as patents, secret processes, copyrights, but it tends constantly to work out an equalization of reward to owners of capital instruments proportionate to the cost of these instruments. So also among enterprisers. Access to this position of peculiar independence in the labor world is thrown open through society's educational system to all who possess the requisite ability. And the enterprisers vie with each other for the possession of managerial methods and of lines of business undertaking which promise unusually large rewards of income.

In this respect land differs from the other productive factors. The owner of an acre of ground in the heart of Manhattan receives an income many thousandfold as great as the income of the owner of an acre outside the city limits. This is not due to any difference in productive service rendered by these two men; for we have seen that economic rent is an income in excess of compensation for all such services. The differences lies in the land itself; in this case it is due to differences in the relative efficiency of situation in relation to the needs of the urban population. All these differences in rent would disappear if man could increase the supplies of natural agents. If plots of the most desirable urban land, mines as productive as the richest in use, acres of the most fertile agricultural ground could be multiplied, none but the best need be used. Not only would differences in rent be wiped out under such conditions, but this entire category of income would disappear; for no natural agent would be exploited beyond the point of diminishing returns. Within limits this is what happens in the case of the other factors of production, and it tends, as we have seen, to remove any differential advantage which one group of workers or capitalists or enterprisers possesses as compared with others. But since this levelling tendency is not applicable in the case of economic rent, competition results rather in increasing the inequalities than in equalizing the incomes of different owners of natural resources.

There is another reason for concluding that land constitutes a

peculiar and unique productive factor. As a general thing all the resources of nature are supplied to man without cost to himself. Man may expend labor in discovering where rich stores of mineral wealth lie hidden and further labor in making them available for human use, but he has not yet learned how to increase the world's stock of gold, iron, or copper. Nor, despite trifling exceptions, does he consciously and intentionally create those differences in situation which are responsible for the superior productivity of certain plots of city land. Labor of generations has improved, it is true, the fertility of some sections of agricultural land, but even in this department of our extractive industry we rely in the main upon the bounty of nature. Being freely provided by nature, the lands, mines, forests, and sources of natural power could be put to the service of man without setting aside for them a share in the social income, provided adequate reward were offered for the labor of discovery, adaptation, and exploitation of these resources. No such statement can be made regarding the services of labor and capital. Labor cannot be drawn into productive service without sacrifice on someone's part; no single unit of our capital equipment can be obtained without the effort of some human being directed to this end. In a society of free men, no labor will be supplied without its compensating wages; little capital will be saved and turned to productive use in the absence of the interest reward; enterprisers will not perform their function if denied the reward of profits. Interest, wages, and profits are necessary forms of income; they are unavoidable payments for the factors of labor, capital, and management — the conditions on which alone supplies of these factors will be forthcoming in the service of society. Even in a Utopia the worker must be paid at least enough to keep him alive and to replace his numbers, present income must be sacrificed if the supply of capital instruments is to be increased and maintained, those who direct the productive system must have some inducement. But in no arrangement of the social order will the payment of economic rent create resources in mines, forests, or lands. These natural agents would exist and be available in the absence of this type of income.

Attacks on private receipt of rent. In the modern world there is increasing antagonism to all forms of extreme inequality — inequality of income and of opportunities for self-development. In its broadest manifestation this supplies the psychological foundation of socialism, a movement of many branches and diverse programs and involving the reform of many other institutions than that of land ownership alone. This larger subject we must reserve for later study, but it may be noted at this point that any attack upon inequality puts the recipients of economic rent on the defensive. Private ownership of land, coupled with the institution of inheritance, tends to create and perpetuate inequality among men both in respect of income and of opportunity. Those who first take possession of the most desirable of nature's productive agents and the heirs of these appropriators through numberless generations to come acquire a position of privilege in comparison with all other members of society. Moreover the tendency of this inequality is to grow greater with the passage of time as economic rents rise under man's increasing pressure upon the supporting power of the soil.

Economic rent as a private income is peculiarly hard to justify on grounds consonant with the common man's impulsive sense of justice and fair play. Economic rent is not a payment for any service rendered by the recipient, for it exists only when the recipient's income is in excess of adequate remuneration for all his labor as a worker, all his service as a saver and investor of capital, all his utility to society as an enterpriser. A service, it is true, is rendered in exchange for economic rent, but it is rendered by nature, not by the rent receiver. In other words, society in permitting the existence of economic rent as a private income is making payment for a service which could be had for nothing.

The problem raised most clearly by urban site rents. These questions as to the ethical soundness of economic rent are brought more forcibly to our attention in the case of urban land than in that of any other category of natural agents, with the possible exception of the mines of older countries where this industry has long been established on a stable foundation. For one reason, this

is because urban site rents are relatively large, measured in units of land area, when compared with the economic rent of agricultural land. There is even room to doubt whether the amount of economic rent returned to the average farmer in a country of vast area like the United States is appreciable. Taking the average annual income of the working farmer and carefully subtracting a normal wage for his labor, normal interest on his capital instruments, and a normal reward for his efforts as manager, it is not probable that much would remain in form of a pure surplus.

Large visible rents. There is no such difficulty in discovering the rent of land in any city of considerable size. The basic cause of this difference between urban and agricultural land is the higher concentration of people on the small land areas embraced in our cities. In 1920 54,000,000 people, or 51.4 per cent of the total population of the United States, were living in cities on a land area not exceeding 10,000,000 acres. The remainder of the country, 1,893,000,000 acres in extent, contained less than half of the people. The total value of city real estate was 84 billions of dollars as compared with a total of 60 billions for all rural districts combined. Early in 1925 a small plot of ground on Manhattan Island with a frontage of 100 feet and a depth of 125 feet sold for $2,000,000, or $20,000 per front foot. Each strip of this land 125 feet long and 1 foot wide was worth more than 65 acres of the richest farm land in the Middle West. On October 10, 1923, the *Chicago Tribune* reported the leasing of a plot of ground 99 by 90 feet for $62,787 a year, on a ninety-year lease. Such examples of high rents and high values are familiar incidents in urban life.

Absentee ownership. Another reason why the problem is more clearly presented by urban lands — and one especially applicable to our country — is that absentee ownership is more common in the one case than in the other. The farm land of the United States is to a large extent owned by working farmers. Whatever of economic rent accrues to these owners is intermingled with their earnings from the active productive functions which they perform and not presented to the casual observer as an income from the privilege of ownership. On the other hand it is customary for the

landowners in the larger cities to withdraw from active utilization of the land and to take their toll from the incomes of industries managed, financed, and operated by others. In the larger American cities only a small fraction of the people own their homes, and practically none at all share in the ownership of the particular land used in the business with which they are associated.

Influence of situation obvious. Furthermore it is easier to see that urban site rents are in no wise attributable to any social service rendered by the recipient. Situation is the factor determining the relative rents of city lands to the exclusion of all other factors. In Salt Lake City one investigator found that the value of land (capitalized rent) fell from $1,800 to $75 a front foot in a distance of 300 feet. This is an extreme illustration of a familiar aspect of urban site rents. It is also easy to see that the rise of rent on city lands is due to impersonal social forces and not to the intentional labor of man. In many of the rapidly growing cities of the West are to be found plots of ground which have never been improved by the labor of man. They lie idle in the heart of the city where men suffer from overcrowding and business is handicapped by lack of room. Though no man has made them one degree more fit for human use than they were in their natural state, these idle plots by reason of their situation alone are now of extreme importance to the life of the community. They are being withheld from use until the exigency of his fellow men empowers the owner to exact the maximum reward in exchange for his mere permission that they serve the needs of society. Clearly the mounting economic rents of land in our cities are the work of social forces which the owner is practically impotent either to call into existence or to manipulate. The location of the nation's transportation system, proximity to an efficient harbor, natural endowment for the support of manufacturers, a swelling volume of foreign and domestic commerce; these, with the consequent thickening of the population, are the forces which determine the aggregate amount of economic rent in such cities as New York, Chicago, Detroit, and Philadelphia. In each place the relative amount of economic rent returned by any one plot of ground is the work of trends in the development

of the city's life, of forces similarly social in character, which determine the relation which the site in question bears to the whole complex of urban activity. That some trifling exceptions to this statement exist may be granted. But in the vast majority of cases nature and the play of impersonal social forces create economic rent and control its amount.

Some twenty years ago was published the economic history of Chicago's most valuable quarter acre — the lot on the southwest corner of State and Madison streets. In 1830, when the population of Chicago numbered 50 people, this plot of ground was bought for $20. In 1845, when the city's population had increased to 12,000, the value of the land had risen to $5,000. By 1860, when the population was 109,000, the lot was worth $28,000 or 1,400 times its original value. In 1875 the population of Chicago had increased to 400,000, and the land had risen in value to $92,500. In 1894, with population at 1,500,000, this quarter acre of ground had a value of $1,250,000. Writing in this year, the author of this bit of rent history stated the matter thus: "Six hundred average Illinois farms would not now exchange for this quarter acre of raw prairie land, and nearly three thousand years of labor of one man would be required to buy it. If 500 years before the Christian era, some man had found employment at the equivalent of $1.50 a day; had, like some Wandering Jew, been preserved through all the vicissitudes of the centuries; had been miraculously sustained without expense for any of the necessaries or luxuries of life; had done his work regularly from that day to this, 300 days in the year without losing a day, and had hoarded all his wages, his savings would not yet be enough to buy this quarter acre of prairie land at the mouth of the Chicago River."[1]

The concept of unearned increment. It has become the practice of students of society, especially those with a bent toward reform, to speak of economic rent as the "unearned increment" of land. As the words imply, this name is used most frequently with reference to the tendency of rent to increase as the years pass. But

[1] Quoted in the *Eighth Biennial Report of the Bureau of Labor Statistics of Illinois*, 1894, p. 279.

there is no reason in principle why distinctions should be drawn between different cases of economic rent. Whether rising, falling, or remaining constant, all true economic rent is, in the common meaning of the term, "unearned" by the landowner; it is an addition or increment to his income in excess of adequate payment for every productive service he performs. The unearned increment may be measured either in terms of annual income or of land value, for we have seen that land derives its value only through a discount of its income.[1]

In using the term, unearned increment, social reformers are of course looking at the matter from the standpoint of society as a whole. It is often urged in defense of rent as an earned income that the present recipient invested his savings in the purchase of the land and is therefore receiving only a justifiable return on his investment. For example, assume that someone pays $40,000 for a site which returns a rent of $2,000. Clearly there will be no unearned increment in the income of this individual — at least until the rent rises — for the yearly rent represents only the normal return on his financial outlay. But the former owner is still receiving $2,000 yearly from the investment of his receipts from the sale of the land, and if we agree that part or all of this income was unearned increment before the sale, it is unearned now from the point of view of society as a whole. To make an extreme case of it, suppose all lands change hands so that overnight a new set of owners appears on the scene. It would then be possible to say that no single existing landowner received more than a just return on his investment. None the less, there would exist in society a group of people — the former landowners — able to share in the social income to the same extent as before, and their share would still be a surplus income in excess of all rewards for productive functions which they performed.

The Single Tax movement. For many years there have been attacks upon the institution of private ownership of land and such proposals of reform as would place it beyond the power of individuals to draw a surplus income from the use of natural agents.

[1] See Chapter XIX.

These reform movements grew in prominence during the nineteenth century as a consequence in part of an increasing popular perception of the true nature of economic rent and partly because of their more able leadership. In modern times the most thorough proposal of reform has been presented in the Single Tax program. The name Single Tax has particular reference to the method by means of which it was proposed that the destruction of private property in land be carried out. It was proposed that the state tax all landowners the full amount of the economic rent, thus turning into a common treasury the entire surplus income from the use of natural agents. This income could then be shared among all members of society either through expenditure by the state upon projects of general social utility, such as education, public health, alleviation of distress, or in the form of relief from the burden of other kinds of taxes — tariffs, excises, income taxes, and especially property taxes upon improvements on the land. For the tax on economic rent was to be a " single " tax — the only tax levied by the state — its productivity being assumed sufficient to discharge all public needs for revenue. This feature of the program was emphasized by the vote-catching slogan of the political party which sought to capture the mayoralty of New York City in 1886: " No taxes at all and a pension for everybody." Rigid application of the Single Tax program would destroy all land values along with economic rent as a private income. To avoid the consequence of government ownership and operation of all extractive industries, it is proposed by the Single Taxers that land be left in private possession on long-term leases, prospective tenants competing against each other for the use of land by offering to bear amounts of taxation which measured their respective appraisals of the economic rent.

The Single Tax movement attained national and international prominence under the leadership of its founder, Henry George. His book, *Progress and Poverty*, published in 1879, is still recognized as a masterpiece of style and eloquence. In a few years after its appearance, it had been translated into many foreign languages and sold to millions of readers in this country and abroad. This

was the time when the agitation for land reform in Ireland was at its height, and the doctrines of Henry George won widespread acceptance in England, where societies were founded to promulgate them. During the eighties and nineties political parties appeared in the United States in support of the program of the Single Tax. The movement was particularly strong in New York City, where Henry George twice ran for mayor, on the first occasion winning a respectable second place in a three-cornered contest. But with the turn of the century it dropped out of American politics entirely and is today represented, according to the estimate of the editor of the *Single Tax Review*, by between 25,000 and 50,000 "convinced single taxers who are in the possession of the full vision." After the death of Henry George in 1897, many Single Taxers began to modify their extreme views and to concentrate on the taxation of future increases only. In this modified form the Single Tax program has been applied by many European cities and, on a wider scale, in the famous Lloyd George budget of 1909. More complete application of the program is found in the policy of concentrating taxes on land value alone to the exclusion of improvements on land. Numerous cities in Australasia and Europe, a few in Western Canada, and at least two — Scranton and Pittsburg — in the United States have adopted this policy. Communal ownership of rent derived from land privately operated is practiced by several communities in the United States, called "Enclaves of Single Tax," such as Fairhope in Alabama, Arden in Delaware, Free Acres in New Jersey, and Tohanto in Massachusetts.

The case for the Single Tax. Laying aside the moral and ethical aspects of the matter, the strongest argument in favor of the Single Tax is that, if levied with such precision that only economic rent would be appropriated, this tax would not operate to discourage economic enterprise or to reduce the total fund of income available to society. It would nevertheless provide a large stream of revenue which might be used for the advancement of general social well-being. The reason for this inheres in the nature of rent itself. If rent is truly a surplus income, it can be withdrawn from its recipient without encroaching on those elements

of his income — wages, interest, profits — which induce him to put forth productive effort. Rewarded as fully as before for his services as a laborer, investor, and manager, the rent-receiver would be under no inducement to diminish his industry because he had been deprived of this pure surplus. The supply of goods and services available for the satisfaction of human wants would, therefore, exist in as great plenty after the destruction of rent as a private income as before. This is because, to repeat a statement previously made, the service for which rent is paid, since it is performed by nature, can be had for nothing.

The case against the Single Tax. Against the program of the Single Tax many weighty arguments of a practical character can be marshalled. Certain of these arguments are of a technical nature, applying to the Single Tax in its character as a measure of public finance. It is doubtful, for example, whether such a tax would raise sufficient revenue to permit of the abolition of all other forms of state revenue. But these technical questions are not essential to our present purpose. Any attempt to carve out the full amount of economic rent from the lands, mines, forests, and other natural agents would raise problems so profound and intricate as to surpass the powers of human intelligence. Even in the simplest cases it is impossible to compute with precision what part of the income from a given piece of land is true rent, what part a return on improvements; and in the case of plots of ground on which improvement has proceeded for generations or centuries, this task of demarcating the rent must be given up as insoluble. Zeal of the tax gatherer to grasp the whole loaf would inevitably cause him to cross that shadowy line which separates producer's surplus from necessary return to productive effort and thus to work unintentional injustice to individuals and to derange the delicate industrial mechanism of society as a whole. Moreover the land area of the world has changed hands so often that it is now, in the majority of cases, to be found in the possession of those who have paid full value for the right to receive the rent. As we have seen, the income of these individuals is merely return on the financial outlay caused by their purchase of land. The sale of land has the

effect of funding the rent and dividing it among various people who are at the time not landowners at all. To search out these hidden portions of economic rent is obviously an impossible task.

Something may also be said in positive defense of the institution of private property in land. The lure of producer's surplus has given energy to the pioneer, the prospector, the searcher after concealed resources of nature. It has acted as a stimulus to man's conquest of his physical environment. It acts today as a force directing the possession of the land into hands most capable of using it to advantage, for the opportunity to win this surplus income plays on the acquisitive sense of men. The same stimulus promotes the readjustment of land to new uses which afford prospect of increased rent to the individual and of more efficient productive service to society. Foresight is encouraged, and lands are withheld from premature commitment to a use of immediate profitableness in view of some future occasion for a different use in which rent will be larger.

Even the absence of these positive social advantages of the institution of private property in land, in view of the practical difficulties in the way of appropriating the full economic rent the attempt would probably result in breaking down institutions which it was not intended to affect. To the average man land is on all fours with other types of private property; his mind does not grasp the subtleties of economic reasoning which differentiate natural agents from capital goods of other kinds. A movement to destroy private ownership of land would be accepted as the beginning of a wholesale abolition of the institution of private property. Once under way, this movement would be extended beyond its original sphere through the activities of those whose particular form of property had been abolished, and the process would lead to a thoroughgoing social revolution. However impressive the abstract arguments in favor of such a reform as the Single Tax, the practical difficulties in its way and the portentous character of its indirect effects make this reform extremely dangerous.

Appropriating future rent increments. It is to be noted, however, that these arguments do not apply to proposals to appropriate

only future increments of rent. If such a program were undertaken with caution, with due regard to the fact that future increases have been reflected to a certain extent in present land values, and in not too greedy a spirit, it would work no harm and would tend to mitigate inequality and increase the funds available to the state for expenditure on socially desirable projects. A large fraction of future "unearned increment," in the narrower meaning of that term, could be appropriated by society to the advantage of all of its members. A consciousness of this fact is spreading in the present-day world, and modern municipalities and nations are beginning to lay progressive taxes on increments of private income which result from the rising economic rents of land and natural agents.

CHAPTER XXXVI

THE NATURE OF INTEREST. THE SUPPLY OF LOANS

Land and capital: landowners and capitalists. Our study of economic rent has disclosed the fact that one group of people in society — the landowners — shares in the social income by virtue of the right of ownership. But landowners are not the only such group. There are many forms of wealth, aside from land, used for productive purposes. The modern industrial system is notable for its vast equipment of machines, buildings, transport systems, etc. — durable instruments of production — which are employed for the production of income. These *capital instruments* resemble land in that they are of use only as intermediate steps in the production of consumable goods and services. But they differ sharply from land in that they are not a gift of nature but the product of man's labor. Since they are man-made, capital instruments will not be supplied unless the income from their use at least equals their cost of production. Hence the owners of these instruments will receive a share of society's income at least sufficient to reimburse them for the expenditures of money and of their own labor involved in the production of capital goods.

Production of capital requires time. In the case of every capital instrument a period of time intervenes between the cost involved in the construction of the instrument and the recovery of this cost through the sale of its product. In many cases this period of time is of long duration. A new railroad, for example, is a group of capital instruments whose initial cost lies far back in the past, when the first steps were taken to obtain the raw materials which eventually took the form of its roadbed and rolling stock. Income from its use is realized only after the lapse of many years. With other capital instruments this time element may be insignificant.

The production of a simple tool may require only a very short time; it may begin to yield consumable income almost immediately after the first costs involved in its production are borne. But whether of longer or shorter duration, the time element involved in the supplying of capital instruments is never entirely absent. This is a matter with which the student has already grown familiar through his study of the roundabout or capitalistic mode of production.[1]

Time preference. The unavoidable lapse of time involved in the production of capital goods means that present enjoyments must be foregone in order that capital goods may appear in the future; that is, those who have command of wealth and services must postpone their enjoyment; in other words, there must be waiting.

Now it is an attribute of human nature to be impatient for present enjoyment. As pleasures are pushed farther into the future, we see them on a diminished scale and they make a smaller appeal to us. It is in man's nature to postpone enjoyment only with reluctance and with a sense of sacrifice, and conversely to feel a distinct sense of gain when, having reconciled himself to waiting for a pleasure which has been promised in the future, he is enabled to abbreviate the period of waiting. Although the pleasure is the same in magnitude, with its nearer approach it grows in desirability and is more highly prized. This being true, identical goods when separated by intervals of time do not have identical present desirabilities. Today's goods are in general preferred to future goods of like kind and amount. We shall examine later certain apparent exceptions, but for the present we may accept the general truth that man values present goods more highly than future goods. To this human trait we give the name *time preference.*

Interest the reward for waiting. These twin facts — that capital formation requires waiting and that waiting is disagreeable to mankind in general — introduce a new element into the reward which accrues to the owners of capital. It means that these supplies of capital instruments must bear two different kinds of cost:

[1] See especially Chapters II and XIX.

the ordinary cost of production and the cost of waiting for deferred returns. In general no enterprise which involves sacrifice of present goods in expectation of a future return will be undertaken unless the return of future goods exceeds the present goods sacrificed. He who exchanges present goods for future balances the two against each other in the scales of *present* appraisal. He will not voluntarily engage in the transaction unless the future return appears at the present moment of equal desirability to the present outlay. Since the dollar in hand and the dollar to be received one year hence do not have equal desirabilities when viewed from the present, the two will not be exchanged for each other unless something in the nature of a surplus or premium is first added to the future dollar. This surplus or premium is *interest*, the reward for waiting. It is a net return to the owners of capital, and society cannot obtain an adequate supply of capital instruments unless the suppliers are allowed a share of the social income which includes this net return in excess of the ordinary cost of producing capital.

Interest is like economic rent in that it is a net income in excess of the cost of the instrument used in production. But its dissimilarities to economic rent are more fundamental than its likenesses. Economic rent is due to the fact that land is limited in quantity and cannot be reproduced. Interest is due to the possibility of reproducing capital instruments provided the necessary waiting is paid for. Interest is a payment for a human service; economic rent is payment for a vested privilege.

The universality of interest. We see that interest appears whenever men consider the exchange of present for future satisfactions. But there is no form of activity open to man whose return does not lie in the future. The building of a house today is rewarded by its shelter in the years to come; the effort expended in planting a tree is rewarded by the fruits borne in the future; the preparation of today's dinner is labor endured sometime in advance of its reward; at least a moment's time separates the filling of a pipe with tobacco from the satisfaction enjoyed by the smoker. But, although we cannot change this order of nature, we do have some choice as to the degree of futurity of our income. Effort

expended today may be rewarded in the immediate or in the remote future according to the direction taken by the expenditure. A day's labor may be consumed in tilling the ground in preparation for a harvest some months hence, or in digging an irrigation ditch whose benefits lie farther in the future, or in repairing an automobile in anticipation of a pleasure ride tomorrow. But whether satisfaction is close at hand or in the remote future, the time element is always present, and interest is universal as a force determining the date at which satisfactions shall be seized.

The form and measurement of interest. Interest is present in all activities of man involving a postponement of satisfactions which might be earlier secured. Interest appears whenever concrete capital instruments are loaned or rented, though the payment in such cases is popularly called " rent " rather than " interest." Even when the capital instrument is used by the owner himself, he receives interest, since by foregoing the immediate satisfactions which could be obtained in exchange for his capital he is bearing the burden of waiting.

When interest appears in this guise, its amount may be measured in relation to the concrete units of capital to which it pertains. To illustrate, we may say that the owner of a particular machine receives $50 per year in excess of the production and depreciation costs of the machine; the owner of a house receives a net income of $1,000 per year; the owner of a motor truck receives a net income of five dollars per day, etc. This is analogous to the method we have used in measuring economic rent, in which the rent income was proportioned to concrete units of land.

But the most common appearance of interest is in connection with " loans " of money or other forms of currency. Thus a borrower, receiving $1,000 in money, contracts to repay $1,000 one year hence and in addition to pay $50 interest. Here the reward for waiting stands out most clearly, and popular usage unhesitatingly applies the term " interest."

In the ordinary loan contract interest is measured as a percentage, the ratio of the amount of the interest, in terms of dollars, to the amount of the capital, also in terms of dollars. In the

example above the interest is said to be at the *rate* of five per cent; $\frac{50}{1000} = \frac{5}{100} = 5\%$. So common and convenient is this manner of measuring interest as a rate that it is customary to apply it also to those cases in which interest is ascribable to particular concrete capital goods. Interest is thus expressed as the ratio between the value of the future annual net income and the value of the capital instrument. Thus instead of saying that the owner of each machine receives a net income of $50 per year, the usual statement is that his net income amounts to five, six, or some other per cent of his investment. The amount will be five per cent if the machine producing the $50 net income is assumed to have a value of $1,000, six per cent if the value is $833.33, etc.

Essence of the loan contract. The prominence given to money in the loan contract and in the expression of the rate of interest must not be allowed to betray us into false conclusions regarding the nature of interest. We have learned to look upon money and bank credit as *purchasing power*, whose important social function is to bestow on the possessor power to appropriate goods and services from the common capital fund or income stream of society. It will become clear on reflection that no borrower desires money for its own sake. When the business man borrows it is with the intention of adding to his factory equipment, enlarging his stock of raw materials, setting labor to work upon a productive process; the real reason, in short, is to acquire goods or services in excess of his present means of payment. The man who borrows for immediate consumption has a similar purpose in mind. The money he borrows will be used to finance a pleasure trip, to pay a doctor's bill, to buy food, clothing, or furniture. With him too the real purpose is to acquire certain kinds of wealth or service in excess of his present means of payment. Behind the borrowing of governments lies a similar motive. The Government of the United States obtains from the buyers of its bonds the purchasing power needed to buy munitions, clothing, or food for the prosecution of a war, or the wealth and labor required to build a canal or irrigation dam; the states borrow to build roads; the cities to erect school buildings,

pave the streets, develop sewer or water systems. In all these cases the real wants of the borrower are satisfied only when the money advanced by the lender has been exchanged for wealth or services in certain forms. The effect of the loan from the borrower's standpoint is to increase his present stock of income-yielding goods.

Upon the lender, the transaction exerts an immediate effect which is the opposite of this. He transfers to the borrower a part of his present means to acquire those products of industry and those services of labor which constitute real income. Whether he deposits his purchasing power in a savings bank which uses it as a basis for a loan to a client, or gives it to a friend in exchange for his I.O.U., or transfers it to a corporation as the purchase price of a bond or share of stock, or buys a government certificate of indebtedness is, after all, immaterial from the standpoint of the immediate effect of the transaction upon himself. In any case this typical lender must make sacrifice to his motive of thrift, sacrifice whose amount is measured by the desirability of the goods which he might have enjoyed if he had chosen to spend his money rather than invest it.

It is apparent that, when the loan is repaid at maturity, a reverse transfer of goods takes place between borrower and lender. Viewing the transaction in its entirety, we see that the lender has given up present income in exchange for future income, and that the borrower has made the opposite kind of exchange, acquiring present income at a sacrifice of future.

Interest not caused by the use of money. This discussion of the nature of the loan contract may appear to be an elaboration of the obvious, but at least it prepares us to perceive that one explanation of the existence of interest is erroneous. Interest is not caused by the use of money as a medium of exchange. The effect of the loan upon borrower and lender would not be different if they were to conduct their transactions in terms of concrete wealth and cease using money as a medium of exchange. No doubt there would be less borrowing and lending in the economy of barter, since the borrower's difficulty in obtaining the kind of goods required for his

purpose would be immensely increased, as would also the lender's difficulty in finding a borrower who could agree to return an acceptable kind of good in the future. But these difficulties would, after all, be matters of detail rather than of principle. Since the effect of the loan, even when money is used as a medium, is to exchange present goods for future, the conditions which cause the existence of interest would remain if this exchange were made directly.

The loan market and the rate of interest. Discovery of the forces which determine the rate of interest is one of the most important problems of economics. Interest being the payment for the service of waiting, the rate of interest is the *price* of waiting; thus when the rate of interest is five per cent per annum, this means that five dollars is the price of waiting one year for one hundred dollars. Like any other price, the rate of interest is determined in the market by the forces of demand and supply. The modern world has evolved an organized market, in which exchanges of present goods for future goods are constantly taking place; *viz.* the loan or investment market. Because of the vast number of transactions which take place in the loan market and the innumerable ramifications of this market through the business world, the rate of interest there determined becomes the governor of the terms on which all types of transactions involving an exchange of present for future income occur.

On the supply side of this market we find the lender or investor. He it is who is willing, for a consideration, to bear the burden of waiting which is indispensable to the creation of capital instruments. On the demand side stands the borrower. He desires present goods and is willing to give in exchange a greater quantity of future goods, thus guaranteeing to the lenders a net income in return for the service of waiting. Demand and supply, thus coming together in the loan market, determine the rate of interest, in accordance with the laws governing prices in general, with which the reader has already become acquainted. Our further inquiry into the problem of the rate of interest will therefore take us into a study of the circumstances affecting respectively the demand and the supply of loans. We begin with the supply side.

The supply of loans. The lenders on the loan market are those who are willing to sacrifice some part of the present enjoyments of wealth which are open to them in return for an addition to their future incomes. Interest is the price which they expect to receive for this service of waiting. And the individual supply curve of the ordinary lender is like the supply curve for any other good, a curve ascending to the right to indicate that the willingness to lend is stimulated by a high interest rate.

But the supply of loans exhibits certain peculiarities not present in the supply of goods in general. There is no doubt that many persons would postpone the enjoyment of some part of their wealth even though they received no interest. Men will make provision for old age and retirement from active business, for the care of their children and other dependents, for unpredictable contingencies, such as accident, sickness, or premature death. Taking advantage of the facilities of the modern loan market, such savers are able to receive interest by depositing in the savings bank, loaning on real estate, buying government bonds or the stocks or bonds of corporations, etc. But if these facilities were not open to them, much of this sort of waiting would still be undertaken without interest. Even as it is, some such free waiting occurs. Many people among the well-to-do habitually carry checking accounts at the banks in excess of their needs, foregoing the gain of interest on the excess for the sake of the sense of security and ease which the large bank balance gives them. These balances, as we know, are made the basis of loans to enterprisers, thus resulting in the formation of capital without interest reward to those who do the saving.

We must go a step further and note that, if this were necessary, some persons would actually be willing to pay something for the safe keeping and later return of their savings. The earliest savings banks in England took fees from their depositors for the service of safeguarding the money and other valuables left in their care. A certain amount of waiting would thus be undertaken even if the savers had to pay for the privilege.

In visualizing the supply curve of the ordinary conservative individual we must therefore recognize these peculiarities. We do

this by commencing the curve below the base line, as in Figure 2, thus indicating that a certain amount of loans would be offered even though there were a charge for the privilege of lending; *i.e.*, at a negative rate of interest, and that a certain larger amount would be offered at zero interest (see Figure 2).

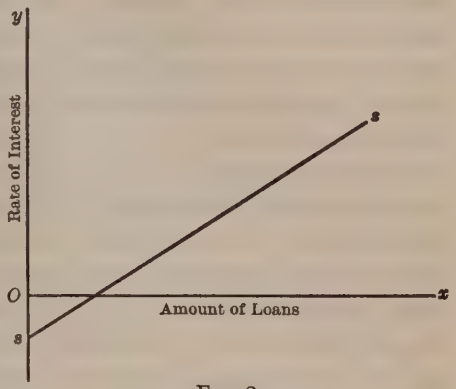

Fig. 2.

Marginal utility in present and future. If the facts of the paragraph immediately preceding appear to contradict our earlier conclusions upon the subject of time preference, the difficulty will be resolved by taking proper account of the marginal character of all satisfactions. When we picture a person weighing the relative desirability of a given enjoyment today and the same enjoyment at some future time, we must not fail to recognize that his comparison is between the respective marginal utilities of the given satisfactions at the two dates and that these marginal utilities depend upon the number of the units of such satisfactions at his disposal at the two dates. To pass from the general to the concrete, suppose that a person, whom we shall call A, enjoys an income today at the rate of $200 a month and is assured of the same rate of income a year hence. If we assume further that all his other circumstances which affect his needs will be the same next year as now, we may safely say that the marginal utility next year of one dollar will be the same as is the marginal utility of a dollar today. Here we have real equality of the present and future magnitudes under comparison, and application of our principle of time preference will tell us that this person will certainly value a present dollar higher than a dollar due one year hence.

Suppose now that another person, B, whose present circumstances are the same as A's, is assured of an income commencing

next year at the rate of $2,000 a month. When he makes comparison of present and future incomes, he compares a present dollar of relatively high utility (one out of 200) with a future dollar of lower utility (one out of 2,000). His personality and judgment may be similar to A's, but his preference of present over future dollars will be vastly higher, since the marginal utilities which he compares are not the same.

Finally take the case of C, whose circumstances are the same as those of A and B except that his present income is at the rate of $2,000 a month, while his future income will be only at the rate of $200. To him present dollars have a marginal utility far less than future dollars. His nature is no different from the others; that is, of equal satisfactions, he values the present higher than the future. But when he compares a present dollar with a dollar due one year hence, he is not comparing equal satisfactions. He will gladly give up a dollar now for a dollar a year hence; he would even be willing to take less than a dollar in the future.

This last example is illustrative of those persons who habitually save something out of their present incomes in order to make provision for future needs, whose saving is not induced primarily by the lure of interest, who would in fact save even if they received no interest or would even pay something for the privilege of safe saving if they had to.

Time preference again. All these cases may be harmonized under a single principle. All normal persons regularly esteem a present satisfaction higher than an equal satisfaction in the future. A given quantity of wealth or income, however, may not represent the same satisfaction in present and future. Depending on their respective personal circumstances, the estimate placed by different persons upon the present quantity of wealth, therefore, may be equal to or greater than or less than the estimate placed upon the same quantity of future wealth. *The difference in the estimate placed upon present and future goods of the same magnitude is called time preference.* Each person has his own time preference. For the vast majority the present good is valued higher than the equal future good, and we therefore call this *positive* time preference.

When one ranks the present good lower than the future good, his time preference is said to be *negative.*

The supply curve. Like supply in general, the curve of total supply of loans on the market is the sum of the individual curves of all potential lenders. No special problem arises in its consideration except to recognize those peculiarities which have already been considered and which cause the curve of total supply of loans to commence below the base axis, thus giving recognition to the

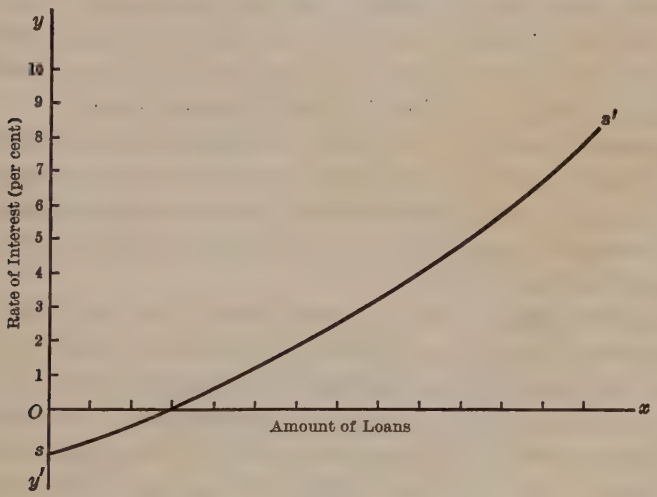

FIG. 3.

principle that a certain quantity of loanable funds would be forthcoming even though the rate of interest were zero or negative. The amount of loans thus obtainable is however relatively small, and the main part of the curve is therefore above the base line, indicating a positive rate of interest. The following curve (Figure 3), which we may assume has been plotted from a hypothetical supply schedule for loans, will serve to illustrate the forces operating on this side of the loan market.

Character of individual saving. The reader should not be led by the somewhat abstract character of the foregoing analysis, to the conclusion that all saving in modern society is done by individuals in a spirit of cold calculation. We know that men are not

actuated by purely economic motives in any department of their lives. It is perhaps true, in respect of the saving done by the average individual, that his action is determined by custom, habit, emotion, more than by meticulous balancing of sacrifice against gain, of the desirability of present against that of future income. The great mass of conservative, cautious, middle class people save, as they say, "all they can." However, since the rate of interest cannot be wholly without its influence as an inducement to save and lend, and since few persons, if any, will be induced to lend more by a low rate than by a high rate, recognition of these facts does not affect the validity of our conclusions as to the general nature of the supply of loans and the shape of the supply curve.

Involuntary saving. The reader should also have in mind that all saving in modern society is not done by the voluntary action of individuals. Certain developments in our economic organization are tending to compel men to save, either against their will or without their knowledge.

Corporate saving. One important tendency of this order is an outgrowth of the corporate form of organization, especially in the large scale enterprises so common in our modern business world. Immediate control of the policy of a corporation lies in the hands of the directors, who are in turn responsible for their actions to the stockholders who elect them. When the corporation has grown to immense size and its stock has become widely scattered among a large number of small investors whose knowledge of the business is vague, the directors are virtually given a free hand in shaping the policy of the undertaking. One of the problems which they decide relates to the disposition of the corporation's net income. This belongs to the stockholders; it constitutes interest and profits upon advances made by them. But it is never the practice of prudent directors to distribute the net income in the form of dividends as rapidly as it accrues. A part is usually set aside as a reserve against future contingencies which may imperil the corporation's ability to pay its debts; another part may be devoted to improving and expanding the material equipment of the business; still another, to experimentation with new processes and selling

methods. In thus disposing of the corporation's income, the directors make deductions from the immediate incomes of the stockholders, more or less without consulting their wishes in the matter, and add to the capital fund of society.

The incentive to accumulate capital in this way is very great in the case of those large corporations which begin their lives burdened with "watered stock." One of the first duties of the directors, if they are to concern themselves with the future stability of their business, is to "squeeze the water out of the stock." The history of the United States Steel Corporation, the International Mercantile Marine, the great railroad systems, and others of our giant corporations is full of illustrations. An examination of the capital accounts of our large corporations over a period of years will bring to light immense annual additions to our capital equipment resulting from this practice of deducting investment funds from the incomes of the people in advance of the receipt of the income by its owners. Certain striking examples may be cited. The United States Steel Corporation began operations in 1901 with capital whose par value was $1,192,146,703, approximately $300,000,000 of which may justly be called "water." By 1922 this corporation had accumulated a surplus out of earnings which amounted to $784,633,838, having also distributed over a billion dollars as dividends. The Standard Oil Company of New Jersey had in 1912 a capital whose par value amounted to $100,000,000. During the next ten years, in addition to declaring dividends of twenty per cent per year, this company accumulated a surplus of $592,021,623. Between 1910 and 1920, according to the estimates of one statistician, incorporated enterprises in the United States reinvested on the average two billion dollars a year. It is of course impossible to say what proportion of these savings would have been made in any case, if the income of the corporations had been distributed as dividends. But there can be little doubt that the amount saved would have been less — perhaps very much less. The perennial inclination of minority stockholders to press for larger cash dividends and oppose proposals to reinvest earnings is evidence of this.

Saving by the insurance companies. Similar significance attaches to modern developments in the field of life insurance. The "level premium policy" — the customary form in which life insurance is sold — results in the accumulation of immense funds in the hands of the insurance companies as reserves against their liabilities to the policy-holders. The entire amount of these reserves, with the exception of accumulated interest, consists of payments made by multitudes of policy-holders out of their respective incomes. They are invested by the insurance companies in the less speculative kinds of investment securities, thus resulting in capital formation. There is little reason to believe that the average policy-holder adjusts the amount of his contribution to the reserves of insurance companies by scrutinizing the rate of interest. Furthermore the practice of making the carrying of insurance compulsory, deducting the premium from the income of the policy-holder without his knowledge or consent — as for example in the case of legally enforced accident, sickness, unemployment, and other forms of social insurance — is spreading among industrial nations. This movement is tending to increase the proportion of society's investment fund whose accumulation bears no direct relation to the interest rate. Of similar import is the practice of leaving estates in the hands of trustees or establishing trust funds for the benefit of the minor children or other relatives of wealthy men. In many such cases, the trustees are bound by the terms of the agreement to reinvest a portion of the annual income regardless of the wishes of the beneficiary.

Other examples could be cited. Loanable funds arising from such sources as these are added at any given time to those resulting from voluntary choice of individuals and tend to increase the amount offered at various rates of interest. Their general tendency is, obviously, to push the supply curve further to the right than it would otherwise be, but not to modify its general shape.

CHAPTER XXXVII

THE DEMAND FOR LOANS. THE RATE OF INTEREST. THE SOCIAL PROBLEM OF INTEREST

The demand for loans arises in general from two distinct groups of borrowers. The first consists of those who *borrow for consumption;* that is, for the sole purpose of increasing their present command of consumable goods at the expense of their future incomes. The second group is made up of business men, whose borrowing is associated with productive enterprise and may be called *borrowing for production.* This group is by far the more important of the two and is the one which exercises the predominant influence upon the demand side of the loan market.

Borrowing for production. Unlike the borrowings of the spendthrift, the business man's loans are expected "to pay for themselves" and in addition to confer an economic gain upon the borrower. It is the purpose of the business man to exchange the borrowed purchasing power for certain forms of wealth and services to be actively employed in his enterprise — a new set of machines, or an additional building, or the paraphernalia of an advertising campaign. If we follow any one of these possible lines of expenditure to its conclusion, we discern that the loan has been employed by the enterpriser to divert a portion of the productive energy of society from the direct into the roundabout process of production. If additions to plant are constructed under the enterpriser's own management, a portion of the borrowed money is transferred to laborers in the form of wages on which to subsist while at work on a non-consumable product; another portion, expended in the purchase of materials, goes to reimburse someone else for a similar transfer of purchasing power to laborers. If these capital instruments are bought ready-made, the expenditure of the borrowed

money gives support to those industries in which labor and capital are devoted to improving the preliminary stages of the process of production. Essentially similar is the result of a loan when it is used to buy intangible services rather than concrete articles of wealth. The enterpriser may have borrowed to pay the salaries of advertising agents, salesmen, or efficiency engineers, or the fees required to obtain a charter or franchise from the State, or the expenses of a promoter who organizes a new industry or of the attorneys who draft articles of incorporation. All funds borrowed for productive purposes are ultimately resolved into advances to laborers.

The productivity of capital. The reader will recall that labor applied in the roundabout way is generally more productive of future income than labor applied to direct production. The capitalistic process, to be sure, is not necessarily more productive; labor may be wasted in over-elaboration of the intermediate stages of production. But the possibility of net gain from the capitalistic process exists, and this possibility is increased by the dynamic character of our economic order. The borrowing of enterprisers rests on this fact of the productivity of capital.

But the enterpriser will ask something more of his borrowed capital than that it increase the physical product of his enterprise. He is not consciously concerned with the broad social aspect of the use of capital; what he is after is a money income for himself, to be derived from the sale of the products of his enterprise. Goods sell on the market at prices determined by the forces of demand and supply. Efficient capital equipment is capable of increasing the quantity of goods produced, but it does not control the prices which they will bring when sold. The business man will borrow only if he expects to sell the increased product of his enterprise at prices high enough to enable him to repay with interest the money expended in the purchase or construction of his capital equipment.

Diminishing productivity from capital. Now the productivity of capital conforms to a law of diminishing productivity. This is found to be equally true whether viewed from the standpoint of society as a whole or from that of the individual business man. In taking

the social point of view, we must define the productiveness of capital as its power to increase the output of goods per unit of labor expended. If only a small quantity of capital is available for the use of society, it will be employed to equip men with the most essential labor-saving devices — for example, the implements used to cultivate the soil, the wheeled vehicles, the tools of the craftsman — whose possession redeems modern man from the abject poverty of primitive races. As the supply of capital increases, other things remaining constant, it will be applied to uses which grow less and less vital to the material welfare of mankind; until the last instalments of the capital supply of a rich society take the form of insignificant improvements on a highly complex machine equipment — effective, no doubt, as labor-saving devices, but in no wise comparable in effectiveness with the first instalments of capital.

To the individual enterpriser all this is pertinent, and in addition he must recognize that the price of his product will fall as the quantity put on the market expands. Whatever the form of capital instrument or productive service to which he devotes his borrowed purchasing power, he will find each one subject to the law of diminishing productivity expressed in terms of money income. Moreover, after he has devised that combination of productive elements which results in the greatest efficiency, to continue expanding the size of this balanced organization will eventually result in diminished money returns from the entire enterprise.

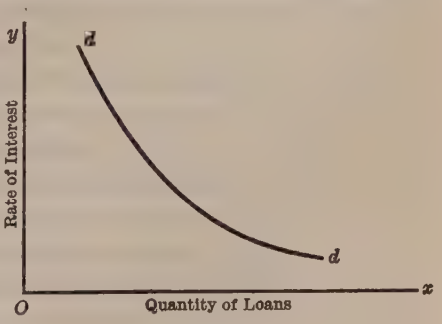

Fig. 4.

From the point of view of the business man, borrowed funds are thus subject to the general law of diminishing marginal utility which we have seen controls individual demand for all kinds of goods. The individual demand curve of the enterpriser-borrower conforms therefore to the shape of the normal demand curve, as in Figure 4, indicating that the amount that

would be borrowed, at any given time, varies inversely with the rate of interest and is such that the marginal productivity of the capital thus obtained is equal to the rate of interest paid.[1]

Borrowing for consumption. There are those whose present needs appear so urgent that they will seek to obtain the present means of satisfaction at the expense of their future incomes and will be willing to pay a premium for the exchange. Such borrowing is for consumption. Repayment of the loan and payment of interest must come, not from the product of capitalistic industry but out of the borrower's future income. There was a time, before the rise of modern capitalistic industry, when most borrowing was of this sort. Today such borrowing is relatively insignificant, not because of any decrease in its absolute amount, but because of the preponderance of the other type of borrowing.

Consumption loans include the borrowing of those who resort to the pawn shops or buy goods on the instalment plan. Certain types of banks, for example the Morris plan banks, cater to such borrowing. Consumption borrowing is by no means limited to the very poor. Well-to-do persons, enjoying moderate present incomes but anticipating larger incomes in the future, are inclined to borrow for present enjoyment. Anyone whose income is irregular may find it to his advantage to finance his personal expenditures by occasional short time loans. Most of the borrowing of political bodies is for purposes of consumption, though government loans for productive investment are not unknown.

Individual time-preference. In the case of productive borrowing, the demand of any business man for loans is determined by the marginal productivity of capital, but when dealing with consumption loans we shall find that the question whether to borrow or not and how much to borrow, at any given rate of interest, is determined by the individuals' marginal rate of time-preference.

Time-preference is a personal characteristic which exhibits wide

[1] The importance attached in this discussion to the productivity of capital should not be allowed to mislead the reader into the notion that this constitutes an explanation of the existence of interest. As previously developed, it is time preference which, by limiting the supply of present goods available for loans, creates and perpetuates the phenomenon of interest.

variations in accordance with individual character and circumstances. Some men are by nature impatient and impulsive, with an undeveloped faculty of foresight or so little self-control that their present desires are unrestrained by the need of providing for future emergencies. With some selfishness or vanity outweighs love of posterity and duty to dependents. Others lack social ambition and have no desire to practice the thrift and abstinence which are essential preliminary steps toward future advance in the social scale for themselves and their dependents.

Other factors lying outside the character of man influence his rate of time-preference. His age and situation in life (especially the number of his dependents) have their effect. The young man before marriage usually shows a higher rate than after marriage; still later in life, when his children are self-supporting and he looks forward to the shortened period of enjoyment left to him, he will normally disclose again a relatively high rate of impatience. Another important influence is the size of the expected income. Other things — personal traits, age, family, etc. — being equal, people with large and stable incomes will have lower rates of time-preference than those whose incomes are small and stable. The latter class feel the urge of unsatisfied present desires. When present income is sufficient merely to sustain a comfortless existence, future needs do not weigh heavily in the balance against the desire for a larger present happiness. Hence the poor are apt to have high rates of time-preference.

Again, among incomes of the same size, differences in the distribution of the incomes over a period of time cause different rates of time-preference. The man whose income, though relatively small in the present, will grow larger as the years pass and who therefore looks forward to a future of comparative prosperity will tend, unless deterred by other factors already considered, to have a high rate of time-preference. On the other hand, the man with a decreasing income is apt to be more concerned with misgivings about his future welfare than with a desire for larger present enjoyment. This phase of his economic condition will tend to give him a relatively low rate of time-preference.

Marginal time-preference. Borrowing for consumption has the effect of enriching the present and impoverishing the future. In accordance with the general law of diminishing marginal utility, each unit of borrowing adds a progressively smaller quantum of satisfaction to the present income and entails a progressively larger future sacrifice. The marginal utility of the present gain decreases, while the marginal utility of the future income sacrificed increases. It follows that the individual's rate of time-preference, based as it is on the comparison of these two marginal utilities, tends to decline as he borrows.

For this reason the amount borrowed depends on the rate of interest; at any given time the amount that an individual would borrow varies inversely with the rate of interest and will be such that his marginal rate of time-preference is equal to the rate of interest he must pay. His demand curve, like that of the enterpriser-borrower, is of the normal shape.

Borrowing of governments. The borrowing of political bodies is essentially similar to the borrowing of individuals for consumption, although some government loans are of the productive sort. In the typical case, the government decides on grounds of social or political expediency to embark on a line of action from which no future financial returns can be expected. The deficit resulting from this expenditure is then covered by the sale of government bonds or other evidences of indebtedness. There is also much short term borrowing in anticipation of receipts from taxes whose income is unevenly distributed throughout the year. The government, like the private individual who borrows to obtain satisfactions in excess of his present means of payment, is in such cases merely anticipating future revenues in order to finance large present expenditures.

But, though these government loans are similar in nature to the consumption loans of private persons, their amount is not determined by the same forces which control the action of the individual with reference to the loan market. The amount of loan funds demanded by political bodies does not vary in any predictable way with the interest rate. A high rate of interest will sometimes cause the postponement of some government enterprise for which borrow-

ing would be necessary; it is more likely to affect the tenure of the loan: *i.e.*, whether it runs for a long or a short term. When a nation is actually at war, the statesmen in charge of the nation's destinies simply borrow all that is needed to bridge the gap between tax receipts and war expenditures, or all that the citizens can be induced to lend for this purpose, paying the lowest rate of interest practicable.

Total demand for loans. The demand of enterpriser-borrowers and that of individuals who desire to increase their present enjoyable incomes may be thrown into the form of schedules and combined to form a total demand. The general rule governing both of these demands is that the amount of the loan varies inversely with the interest rate. Government borrowing is not easily reduced to a demand schedule because non-economic motives weigh heavily in all political activity. The government comes into the market as occasion requires, with a predetermined demand for loans, and offers the lowest possible rate of interest. Though there are times when the necessitous borrowing of governments is controlling, as a rule the demand of enterprisers so vastly outweighs all other types of borrowing that it governs the market. Borrowers for consumption and governments in peace times must adjust their demands to a loan market dominated by business men whose borrowing is based on the productivity of capital. Strictly speaking, however, all three types of demand must be added together to give a measure of the forces operating at any given time on the demand side of the market.

Determination of the market rate of interest. As already noted, the rate of interest is a price — the price of waiting — determined, like other prices, by the forces of demands and supply acting in the loan market. Such peculiarities as attach to this particular price have now been disclosed by our study of the demand and the supply of loans. It remains only to conclude this part of our study by stating that in any given market at any given time the rate of interest and the amount of loans is determined by demand and supply, as illustrated in Figure 5, in harmony with the general laws of price as set forth in Chapter XIII of this book.

Relations of the individual to the loan market. The rate of interest, like any other price, is the effect of the activities of a multitude of individuals demanding or supplying loan funds. But, again like other prices, this rate is also one of the causes of the

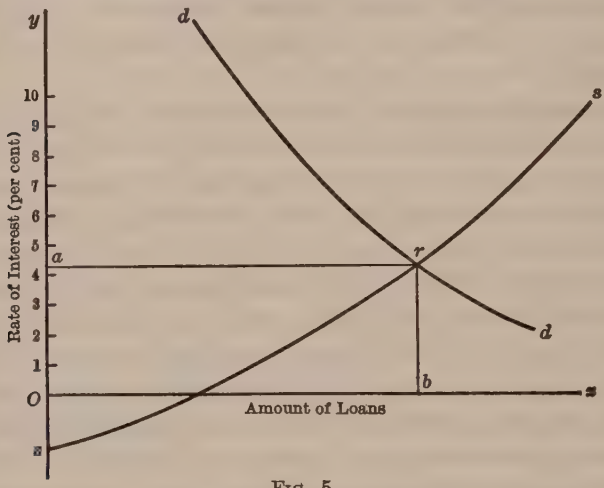

FIG. 5.

amount offered or taken by any given person. The individual looks upon the existing rate of interest (as he does upon any other price) as fixed by forces beyond his control; he is concerned with adjusting his activities so as to make the best of conditions as he finds them. Certain principles controlling these individual adjustments follow readily from our previous study of the demand and supply of loans.

(1) *The individual lender.* A man will lend when his rate of time-preference is lower than the rate of interest. He will continue to lend until his marginal rate of time-preference is equal to the interest rate. Theoretically he gains from all units of his loan except the last and, by lending up to this point, makes his gain a maximum. In a perfect loan market, the result of the activities of the individuals will be to give them marginal rates of time-preference equal to each other.

(2) *The enterpriser-borrower.* This person borrows when the marginal productivity of capital under his management exceeds the

interest rate, and he continues to borrow until this discrepancy is removed. Since it is the income from the marginal unit which equates with the interest rate, there is a gain to be derived from the use of borrowed funds for productive purposes. This surplus is not retained by the enterpriser-borrower. It represents that part of the income of the enterprise which is not absorbed by the interest cost, and it is distributed among the other productive factors which comprise the business unit. The tendency for each enterpriser to borrow until the marginal income from the loan equals the interest rate results, in a perfect loan market, in distributing the capital supply of society among the different business units and the different forms of capital instruments so that its marginal productivity is everywhere the same. This tendency works toward the end that society receives the maximum product from a given capital equipment.

(3) *Consumption loans.* An individual will borrow for consumption when his time-preference rate is higher than the rate of interest, and he will continue to borrow until this discrepancy is removed. Like the buyer of goods, he gains utility from all units of his loan except the last one. In a perfect loan market the tendency is for the rates of time-preference of all such borrowers to reach equality. As stated above, however, the borrowing of governments is not generally controlled by this principle.

(4) *The alternative of lending or borrowing.* The capitalist is generally in a position to be either a lender or a borrower, his choice depending upon the relation of his own rate of time-preference to the market rate of interest. Let us suppose that a certain person's rate of time-preference is six per cent. If he finds that he can borrow at five per cent, it will be to his advantage to do so. But if the market rate of interest is seven per cent, he will have no inducement to borrow. On the contrary, if he is the possessor of any considerable income, he will find it to his advantage to lend some part of his present resources, up to the point at which his own rate of time-preference becomes seven per cent. In particular the business man who is himself an owner of capital will enlarge his enterprise with borrowed funds if the market rate of interest is

lower than the productivity of his business, whereas he will contract his operations and lend his free capital to others when the opposite relation holds. This possibility of shifting from one side of the market to the other contributes much to the sensitiveness of the loan market.

The rate of interest and the quantity of money. We have previously had occasion to refute the popular notion that interest is caused by the use of money as a medium of exchange. We are now in position to dispose of another stubborn money fallacy which will not down, despite the repeated disproofs of economic science. Each generation brings to the front a group of reformers who, either themselves debtors or impelled by a compassion for the debtor class, propose to abolish interest or to lower the rate of interest by increasing the quantity of currency. Their program is based on the belief that the rate of interest varies inversely with the quantity of money, this belief resting in its turn on the assumption that borrowing constitutes a demand for money and that therefore the price which the borrowers pay (*i.e.*, the rate of interest) must be controlled by the relation of this demand to the supply of money. This assumption would not be far from the truth if the term " supply " were used in a sense coördinate with the term " borrower's demand "; if, that is to say, it referred to the supply of money available for loans. But we have seen that the borrower does not want money itself, but money's worth. Since the effect of an increase in the quantity of money is to reduce its purchasing power, a larger amount will be required to accomplish the purpose of the borrower. It follows that, as the price level rises in consequence of an increasing supply of money, the demand for loans measured in dollars and cents also rises, thus tending to offset the effect of the increased supply.

Experience seems to bear out the opinion that the interest rate fluctuates with the amount of money in the country. The transactions of borrowers are carried on to a large extent with the banks, and their knowledge of the loan market is formed out of their experience with these lending institutions. From this experience they learn that when the quantity of money in the bank's vaults

rises or falls in proportion to loans carried by the bank, the interest rate on loans tends to move in the opposite direction. The banker will explain the low rate of interest which accompanies a relatively large cash reserve by the statement: "money is plentiful," and the high rate which accompanies the relatively small cash reserve by the statement: "money is scarce." But here we have to do, not with the amount of money in the country, but with the amount in the banks and with the motives which induce people to deposit their money as an alternative to spending it. If the bank reserves increase only because of an increase in the total stock of money — that is, if the proportion between the amount of money in the banks and that in circulation remains the same — the greater lending power of the banks will be offset by the increased amount of money required by the borrowers to conduct their transactions on the higher price level. If however the bank reserves increase faster than the amount of money in the country, this is an indication that the people, for one reason or another, are increasing the proportion of money available for loans. In this case there will be a tendency for the interest rate to fall, but obviously the cause of the fall is to be sought in the motives which have induced the people to save more and spend less, not in the amount of money in the country at large. Nor can it be shown that an increase in the total quantity of money will enable the people to save proportionately more than formerly when account is taken of the rise of the price level which money inflation causes.

Transition effects. We conclude then that the rate of interest is not necessarily either lower or higher when the amount of money is large than when it is small. But this considers ultimate effects only, leaving out of account how the interest rate is affected by the transition phenomena of the period which intervenes between a low and a high price level. We know that when prices are rising under the influence of currency inflation, business men strive to take advantage of the favorable market by expanding their enterprises through increased borrowing. If, under these conditions, the new stocks of currency are so distributed that the bank reserves temporarily fail to maintain their ratio to the amount of money in the

country, the need for loans will outstrip the quantity of money available for loans and the interest rate will rise. On the other hand, when prices are falling because of a contraction of the currency, money may find its way into the banks in unusual amounts, due to the tendency of the people to restrict their expenditure during periods of depression. At the same time, the general feeling of business uncertainty will tend to reduce the demand for loans; hence the interest rate may fall when the quantity of currency is decreasing. The effect on the interest rate, in other words, will be the opposite of that expected by those who believe that the rate of interest varies inversely with the quantity of money.

We should note also that the effect of a rising price level is to reduce the real worth to the lender of the principal of loans which he made prior to the rise of prices. A loan of a thousand dollars loses one half of its purchasing power if the price level doubles before the loan is repaid. Rising prices therefore penalize lenders in the case of all pre-existing loan contracts and benefit the borrowers to an equivalent extent. Conversely, a falling price level impoverishes borrowers and enriches lenders. It is to be expected that borrowers and lenders will take account of these effects of a fluctuating price level and attempt to guard against them by raising or lowering the interest rate. During a period of rising prices the interest rate tends to rise as an offset to the depreciation of the principal of the loan; during a period of falling prices it tends to fall as an offset to the appreciation of the principal. This tendency coöperates with the one just discussed to cause a rising interest rate when the supply of currency is increasing, and *vice versa*. These, to repeat, are transition phenomena. They disappear when the price level is again stabilized.

Interest rate a factor in all valuations of income. Our study of interest has shown that it arises in connection with all forms of human activity which involve an exchange of present income for future. In explaining the forces which determine the rate of interest we have chosen to center attention on loan contracts because they serve as a concrete example of this type of transaction and because a highly perfected loan market has developed in which the

interest rate is established as an explicit price. The market for loans and investments holds such a dominating position in modern industrial nations that the price (interest rate) established there tends to control the activities of the entire business world in so far as these activities involve a linking of present and future values. This is because the loan market presents, to all members of society who are disposed to surrender present income in consideration of a future return and to all who are disposed to make the reverse exchange, an opportunity to carry out the transaction at a predetermined rate of interest.

Aside from loans and investments, the most commonplace type of transaction involving an interest rate is the evaluation of articles of wealth on the basis of their future incomes. The reader has already been reminded that the return for all human activity lies in the future, more or less remote, and he is equally aware that the value of every article of wealth depends upon its expected future income. The buyer of an article of wealth is exchanging present purchasing power for future satisfactions, and the discount process enters into the determination of all prices of concrete goods. The reader is referred to Chapter XIX, where these principles have been fully discussed.

The interest rate is of peculiar significance in the valuation of land. Capital goods in general are made by man. If the discounting of their income results in a price which is above the marginal cost of production, the suppliers will increase the quantity offered for sale. This in turn will reduce the income from the marginal unit and cause the price to fall until it is equal to the marginal cost. The sale price of capital instruments therefore is controlled by the joint influence of marginal cost and discounted income. Land however is fixed in quantity and has no cost of production; hence its value can be determined only by the discount process. As rents increase and the value of land rises in consequence, the impossibility of adding to the quantity of land removes the only effectual limitation upon its sale price. The value of land thus depends solely upon the amount of the expected future economic rent and the rate of interest.

Lack of uniformity in interest rates. We have been speaking of the interest rate as if there were a single uniform price applying to all types of investments in the loan market. But when we turn to the business world for illustration of our theory we are impressed at once by the absence of any such uniform price. Different rates of interest appear in the market at the same time, each rate applying to a particular type of loan transaction. For example, the following schedule of rates was published on June 16, 1925, for transactions in the New York money market:

Type of loan	*Per cent*
Call loans	3¾
60–90 day loans	4
Commercial paper	
Best names	4
Other names	4¼

On the same day long term bonds sold at prices which returned various rates of interest. The following examples will serve to illustrate these variations:

Type of bond	*Interest return on price*
U. S. Treasury, $100, 4%	3.8%
French government, $100, 8%	7.3
United Kingdom, $100, 5½%	5.1
New Haven Railway, $100, 4%	6.8
New York Central, $100, 5%	4.9
Pierce Arrow Company, $100, 8%	7.7

While these variations existed in rates of explicit interest, business ventures financed by their owners were returning rates of implicit interest which also varied widely.

Some of the variations appearing in the first of these tables are due to conditions peculiar to the banking business. As we know, a large part of the banker's liability consists in deposits payable on demand against which he must maintain assets in liquid form. The reader will recall that the bankers in a money centre like New York receive on deposit large amounts of cash funds from banks in inland cities subject to continuous withdrawal and replenishment. For obvious reasons it is not profitable for these bankers to hold such deposits in the form of idle cash. The call loan is the best

form of investment for funds of this kind; it is resolvable into cash on demand and secured against loss by a deposit of collateral. For this reason the call loan in a large money centre commands a peculiarly low rate of interest. This rate also fluctuates sharply in response to variations in the amount of money seeking this peculiar form of investment. On June 16, 1924, for example, the call loan rate stood at 2 per cent, slightly more than half what it was a year later. The necessity of maintaining a large part of his assets in quickly convertible form accounts also for the banker's practice of charging a lower rate of interest for short term loans — 30 to 90 days — than for long term loans of equal security.

True and nominal interest rates. On loan transactions not subject to these exigencies of the banking business, variations in the interest rate are often more apparent than real. We must view interest as a payment for a single kind of service — the service of postponing income by a voluntary exchange of present goods for future. If the man who renders this service performs at the same time a number of other functions and receives a single payment as a recompense for all of them, only a part of the payment can strictly speaking be called interest; the remaining part is composed of as many different elements as there are different functions to be rewarded by the payment. Now it happens that the term "interest" is commonly used in the business world to define just such a composite payment, of which only one element is true interest.

The rate of interest in ordinary loan transactions contains, in addition to interest proper, an element which should be considered insurance against risk. When dealing with true interest we must assume that lenders are assured of their future return. If, in addition to sacrificing present income, they must bear risk of loss, their reward must be increased by a payment sufficient to induce them to take the risk. The differences in the rates of interest earned on the several bonds listed in our table may be attributed chiefly to this element of risk. For our purpose it may be assumed that through all these different types of investment runs a common true interest rate, which is augmented in each case in proportion to the risk involved.

The capitalist who undertakes to manage his own investments must perform other duties in addition to simply postponing the enjoyment of his income. He must expend effort in searching out profitable lines of investment, collecting information relating to the industries they represent, following the trend of market conditions in order to safeguard his interests. In certain types of investment — as, for example, the ownership of rented houses, office buildings, etc. — he must also undertake to collect the interest periodically and to care for the capital. In all such cases true accounting would segregate a portion of the return loosely denominated interest and give to it the name wages of management. The fact that a separate group of men — investment brokers, trustees, etc. — have developed to relieve the lender of some of these duties, in consideration of a stated remuneration, shows that the duties are not inevitably combined with the services of the lender. When the investor assumes these tasks himself, it is proper to consider him entitled, in addition to interest, to the same remuneration that he would have been compelled to pay another to perform them in his behalf.

Wages of management account for an even greater part of apparent interest when the business man uses his capital in his own business and, in his cost accounting, either neglects to allow himself a salary or fixes his salary at a figure below his real worth as a manager.

Yet another element commonly confused with interest is depreciation. When an investment takes the form of capital instruments which are wearing out, a portion of the capital value is returned with the interest each year, thus making it appear larger than it really is. If this continues without the knowledge of the investor, he will find in time that his capital no longer exists; it has been consumed in the process of returning an annual income. The prudent business man will anticipate this result and provide against it by setting aside a certain part of the income in a depreciation account, viewing only the remainder as true interest.

In our study of the rate of interest then it must be borne in mind that we have to do with interest alone, and not with the other

payments which are commonly confused with it. An investment whose nominal and true rates of interest are identical is one in which the investor bears no risk, exercises no managerial powers, and receives no depreciation in connection with the income. Very probably, as a practical matter, no investment can be found which measures up exactly to this ideal. But it is approximated in the case of the long-term bonds of a stable government, whose return may be accepted as a fair exemplar of true interest.

Changes in the rate of interest. When we examine the trend of the interest rate over long run periods we are impressed by its relative stability. The rate, it is true, fluctuates rather sharply in short run periods, showing a tendency to rise in the wake of rising prices during the " prosperity " phase of the business cycle and to fall to a low point during the subsequent reaction. The reasons for these fluctuations during the business cycle have been explained on a preceding page (See Chapter XXV). Other short term fluctuations are due to temporary maladjustments between demand and supply in the loan market. Interest is a sensitive price, which responds promptly to sudden changes in market factors. But all these short term fluctuations are lost in a uniformity of trend over long run periods such as is exhibited by no other price. In Holland in the seventeenth century, when the interest rate fell to three per cent, this rate was only slightly lower than when the Roman Empire was at its height. During the following century the rate rose slightly, only to fall during the nineteenth century to the same low point. The past five decades have witnessed some rise in the rate.

Comparisons between the interest rates of different countries at widely separated points of time are not wholly reliable, because of the impossibility of making accurate allowance for the element of risk. The rate is especially sensitive to conditions of civil disorder and international war. Yet the testimony of history as to the stability of the interest rate is impressive, the more so when one considers the complete revolution in the technique of production, the rapid change in population, the broadening of markets, which have occurred during the centuries. The trend on the whole has been in a downward direction during periods of history when se-

curity of private property has been undisturbed by civil disorder and when the equilibrium of the loan market has been unaffected by heavy military expenditures.

Future of the interest rate. A half century ago this tendency of the interest rate to fall had been supported by so much historical evidence that it was accepted by many economists as a sort of natural law. A declining interest rate, if produced by automatic economic forces, can be accepted as a sign of improving social welfare. It is evidence that the supply of capital is so plentiful that it can be turned to uses where its marginal productivity is low. The influence of a low interest rate on the distribution of income is to diminish the share going to the propertied classes and to increase the prosperity of that great section of the population which shares in the social income through the sale of its labor. This is recognized as a condition contributing to the stability of the social order and to the general improvement of mankind.

Predictions as to the future of the interest rate are extremely hazardous, but it is possible to speak of certain general forces which will control. A continued fall of the rate of interest can only be the outcome of such a general lowering of the rates of time preference of the people as would indicate a relatively higher regard for future income. Many tendencies are at work today to produce this result. Present income for the average man is larger than it has been at any time in the world's history; his present range of satisfactions is broader; a surplus exists above his vital needs. This provides the material basis which makes it possible for the average member of society to postpone the consumption of a portion of his income. Other developments are increasing his disposition to do so. The progress of education and general enlightenment tends to increase the foresight of the common man by clarifying his view of the future and its hazards. To the same cause we may attribute the progressive expansion of human wants of a sort other than the want for momentary physical gratification. An urgent desire for attractive and spacious homes, for good clothes, tasteful furniture, books, travel, and leisure time causes the middle class man to take thought if by abstinence and present

economy he can hope to acquire them. Above all, the increasing democratization of society, removing the barriers of birth and status which formerly confined men within certain rigid social classes, sets free the powerful impulse of social ambition to spur men to aspire to positions of greater dignity and significance in the eyes of their fellows. The dominant ideals of a capitalistic society demand, as a prerequisite to advance in the social scale, the accumulation of wealth. These tendencies are supported by the increase of public order which has followed the establishment of strong governments in the modern world.

On the demand side of the market in modern times, borrowing for investment holds chief place. This was not the case before the development of the capitalistic social order, nor is it now true among backward people. The ratio between the two types of borrowing is occasionally upset by the periodic demands of governments in times of war. But during the intervals of peace investment borrowing is dominant. Over long periods therefore significant changes on the demand side of the loan market result from developments in the industrial world, and these developments are tending consistently in the direction of increasing the amounts of loan funds demanded at given rates of interest. Of such effect have been the inventions of machinery, the adaptation to productive uses of steam and electrical power, the discovery of rich sources of raw materials, the perfection of methods of transportation, the improvement of business organization, which are the outstanding features of social progress since the Industrial Revolution. All devices which reduce the cost of producing goods afford to the enterprisers the prospect of increased money income if they can obtain command of the present means required to exploit them. These expectations impel the business man to borrow more largely at the existing rate of interest until the gain is wiped out by the declining productiveness of the new opportunities. There is no doubt that we shall see a sustained expansion of the demand for investment loans during the years to come. On the other hand, for reasons we have sketched, we may assume with confidence that the supply of loan funds will increase. Leaving out of account the

possibility of civil disorder, social revolution, or international war on a large scale, we conclude therefore, in the phrase of an eminent writer on economics, that the future history of the rate of interest depends upon a "race between accumulation and improvement."

The capitalist class. Whatever may be the outcome of these forces determining the rate of interest, there is no doubt that the rôle played by capital in the productive system of society must grow in importance with the passing years. Our civilization is founded on the use of great masses of capital; progress in human well-being — nay, the mere support of life for an increasing population — requires that we set aside each year an enormous amount of present income for exchange against the income of the future. The capitalistic mode of production calls for two distinct kinds of service: working and waiting. Neither of these services by itself can maintain the flow of consumable income at a volume sufficient to satisfy the wants of men on the present plane of living. Man's labor of hand or brain, unless reinforced by his willingness to wait and to endure the costs of waiting, must be applied to direct production without the help of those product-multiplying aids which we call capital instruments.

Any commonplace incident of modern productive life will illustrate the indispensability of the service of waiting. After the World War the United States, in company with other nations, suffered from a great dearth of houses. Millions of people were crowded into tenements unfit for human habitation; all agreed that one of the most vital needs of the people was for shelter. Let us suppose that a group of people who possessed no resource other than their labor and their intelligence had undertaken to produce this necessary good. Their first step would be that of appropriating from the resources of nature the necessary supplies of lumber, ore, stone, cement, and other building materials. Each of these must then be fashioned by a time-consuming process into forms suitable for the purpose in view. More time must elapse while they are transported to the places where the want for shelter is felt. Then must begin the work of construction looking toward the creation of consumable income. At no stage in this long process would the

laborers be able to subsist on the current product of their labor. If we relied on labor alone the undertaking must be abandoned at the very start, and no houses would ever be built. When therefore some member of society who is not in the group of workers comes forward with an offer to "finance" the undertaking — *i.e.*, to transfer a part of his present income to the workers in exchange for their non-consumable products — must we not conclude that his service of waiting is as indispensable to the production of this vital necessity as the services of the workers? Or, to change the conditions of our illustration, if we assume that the workers are supplied with resources of purchasing power sufficient to support their lives while their product is being completed and that they finance the undertaking themselves, will they not perform a double service for which they are entitled to a two-fold reward?

The existence of explicit interest as a separate share of the social income is merely evidence of a certain kind of division of labor in modern society. As our productive system grows more capitalistic, the numbers of those who draw interest incomes must increase, and the size of the incomes of this type which accrue to the larger capitalists must expand. Although, as we have seen, the capitalists perform an indispensable productive function, this service is not rendered at the time the interest income is received but has been performed in the past when the saving was done. Indeed in many cases the service of waiting is not even performed by the same man who receives the interest but by some predecessor from whom he has inherited his wealth. One consequence therefore of the creation of a distinct capitalist class is the existence of a group of people who appear to enjoy the right of sharing in the income fund of society while making no concurrent contribution to that fund.

Opposition to interest as a private income. The spectacle of the receipt of incomes by individuals who appear to be idlers affronts the sense of justice of many a man who does not reason clearly with regard to the capitalist's function in society. Other grievances associated with the existence of interest as a private income add to his bitterness and give impulse to his attack upon the institution. By inheritance, ownership of capital falls to men and

women who live in opulence as drones in the midst of a busy world. Powers other than the mere right to receive interest are gathered into the hands of the capitalists — power of command over the lives and welfare of propertyless men and women, power to influence for good or ill the functioning of society's industrial structure, preponderance of power in the political world — which give the capitalist a sinister aspect in the somewhat distorted view of the common man. Reaction against the institution of interest as a private income leads to organized reform movements which strike at the very foundation of modern capitalistic society.

Early regulation of interest. Though we cannot undertake at this point in our study to give to these movements the examination which they deserve, their direct bearing upon the subject in hand requires us to take some account of them here. In part the modern attack upon interest is an inheritance from an earlier day and a simpler stage of societal development. Throughout recorded history there have been individuals and communities which have condemned the receipt of interest by lenders as an immoral and indefensible practice and have taken steps to abolish or strictly to regulate it. Aristotle decried the taking of interest on the ground that money was " barren " and could have no legitimate " offspring." The Mosaic Law prohibited the payment of interest between Jews. The leaders of the early Christian church carried over this condemnation and applied it to their own communities. By the laws of Rome interest-taking between Roman citizens was similarly forbidden. Perhaps the best known example of this attitude is supplied by the Church during the Middle Ages. For over a thousand years, indeed down to the thirteenth century, the Church Fathers carried on an unceasing attack upon interest. By the Canon Law its receipt was strictly forbidden to Christians on pain of spiritual punishment in this world and the next, a prohibition which had the incidental effect of throwing the business of money lending into the hands of the Jews, who were considered at all events to stand outside the pale.

These laws became obsolete and gradually disappeared in all modern nations. But vestiges of them remain in the form of

" usury " laws, and the point of view which they represent comes to the front in modern times, not only as a phase of the program of the socialists, but also in the form of sporadic political movements among more conservative men. During the earlier period the term " usury " defined the entire amount of interest received by the lender; all of it was condemned as an immoral income and all prohibited. Today we divide the interest payments into two parts and employ the word usury as a term of condemnation applied to the excess of the interest rate above an amount vaguely fixed by changing ideas of justice and fair play. The usury laws of the American states, which prohibit the taking of interest above a fixed maximum on certain types of loan transactions, are of this character.

There was some ground for moral condemnation of the interest income in early stages of societal development. Almost all loans in that day were loans for consumption. The spendthrift borrowed to enlarge his supply of consumable goods; the necessitous man, the widow, the victim of misfortune were compelled to borrow to eke out a day-to-day existence. The king or the lord of the manor borrowed to cover the expenses of a war or crusade or to heap favors upon the court favorite of the day or bedeck with baubles the ubiquitous mistress. Such loans were indeed " barren," to adopt the expression of the moralists of the time. As long as they remained the typical form of loan transaction it was difficult either to explain on economic grounds why interest existed or to justify its existence on grounds of morals. Certainly such a borrower was not enabled to bear the charge which interest laid against his future by creating income through his use of the borrowed funds, as does the enterpriser-borrower today. And in the common run of cases the position of the creditor smacked of a grinding of the poor, a fattening upon the necessities of his brother man. The usury laws of the time were passed in defense of the weaker man, the debtor, to protect him from being surrendered by misfortune into the hand of the voracious money lender.

The modern field for regulation. Where this type of loan has carried over into modern society we continue, with general ap-

proval, to regulate the interest rate out of concern for the borrower. In the industrial cities of all countries there exists a class of lawful pirates who have acquired the illuminating title of "loan sharks." They flourish among the timid and ignorant and at the same time most necessitous sections of the wage-earning class, taking advantage of the proverbially high time preference of this class to involve them in bonds of indebtedness, exacting exorbitant rates of interest for petty loans advanced usually on the security of chattel mortgages. Familiar to all city dwellers is the ubiquitous pawnshop of such districts — another type of lending agency which caters to an occasional but highly urgent want of present income among people who, for one reason or another, are excluded from the general loan market. In many countries whole sections of the people are prone to fall into the clutches of an inhuman creditor through their borrowing for improvident purposes. The position of the petty farmers of India or Egypt, the villagers of Russia, and the ignorant Turkish peasants of Anatolia has given frequent testimony to the deplorable social consequences which may result from allowing a group with high rates of time preference to fall prey to the loan shark. Though interest is no longer prohibited on loans for consumption, all modern countries make some provision for the protection of the most necessitous borrowers from the worst consequences of their thriftlessness. France has established public pawnshops where small loans may be had at reasonable rates of interest. England has opened quasi-public banks in the rural districts of India and Egypt which underbid the village Shylock both in terms of interest charge and in facilities for discharging the indebtedness. In America the states regulate the pawnshops with laws fixing a maximum interest rate and requiring accessible records of all transactions, while public-spirited citizens have set up a system of loan offices designed to break the hold of the lender on chattel mortgages upon his unfortunate debtor.

Futile and injurious interference with the interest rate. But the typical debtor of today is not the timid victim of misfortune whom imagination pictures as struggling in the grip of a brutal money lender. On the contrary he is the most able, most self-reliant of us

all — the business enterpriser. As modern society emerged from the Middle Ages, the capitalistic system of production began more and more to provide the outlet for the accumulated funds of the savers. Borrowing for production in the normal course of events places the borrower in position to meet the interest charge without trenching on the income which otherwise would accrue to him. However, despite the fact that the day has passed when spendthrift loans comprised the major part of the business of the loan market, the antipathy toward interest — especially a high rate of interest — remains and finds expression in laws intended to restrict the rise of the rate in the financial world. Admitting that exceptional cases arise which require exceptional treatment, we should see that such laws are in general both futile and injurious. They are futile because, when conditions arise which enable the enterpriser-borrower to pay high rates of interest and therefore create in him a disposition to do so if necessary, the purpose of usury laws will be defeated by connivance of both borrower and lender. They are injurious, both from the standpoint of the individual borrower and from that of society as a whole. If any branch of business or any geographical locality is handicapped by lack of capital — and this is invariably the case when complaints are heard about the high rate of interest — it is obviously to the advantage of all enterprisers to facilitate an increase in the supply of loan funds. The basic cause of the condition cannot be removed by arbitrarily relieving present debtors of a part of their interest costs. What is needed is renewed inducement to saving and investment. Usury laws, by depriving savers of a part of their inducement to save and by injecting into the loan market the incalculable element of political interference, tend to perpetuate the evil which lies at the root of the protest — a paucity of loan funds. Not only the individual who is attempting to carry on productive enterprise under the handicap of small capital supply, but the entire community which shares in the benefit of the capitalistic process, is benefited by allowing the interest rate to fluctuate freely in accordance with the forces of demand and supply.

CHAPTER XXXVIII

THE GENERAL LAW OF WAGES

The rewards of human exertion. In the six chapters immediately preceding, our attention has been devoted to those shares of the social product which accrue to the owners of land and capital. In modern society such owners of productive wealth are clearly in the minority; the vast majority of mankind must derive their income from their own exertions; that is, from labor. The term *labor*, as generally used in economics, has a broad content, including specifically the exertions of the members of these three groups: (1) those who sell their labor, whether manual or mental, to employers in return for wages or salaries, to which class the term labor in a narrower sense is sometimes confined, (2) independent professional men and women, such as lawyers, physicians, artists, and technical experts, who render services to their clients in return for fees and professional earnings, and (3) business men, enterprisers, entrepreneurs, who own and direct industrial enterprises and receive their remuneration in the form of profits.

While it is customary to distinguish wages, salaries, fees, and professional earnings, it will be found convenient for economic analysis to lump these forms of income under the single general term, *wages*, which, in that sense, may be defined as *remuneration received from other persons in exchange for labor*. This includes the incomes of all those in the first two labor classes, as distinguished from the profits of the third class. The investigation of the causes which determine wages will occupy us during the next two chapters, following which we shall give our attention to the subject of profits.

The problem of the rate of wages. Since wages result from the sale of labor, the rate of wages is a species of price governed by the forces of demand and supply in the labor market. But a study of these controlling forces presents peculiar difficulties, which make

an explanation of the wage rate one of the most baffling problems in the whole domain of economic science. We search in vain through the business world for a uniform rate of earnings for different workers similar to the rate of interest which, with relatively small and explicable variations, determines the incomes of different capitalists. The great body of the working people is divided into a bewildering number of groups and sub-groups, each enjoying a different rate of income, and within these separate groups the earnings of individuals will be found to vary in such fashion that any attempt to formulate a general law governing the rate of wages appears hopeless at the outset.

In view of these complexities, our first step must be to simplify the problem. This we can do, first, by limiting our attention at the start to some group within the great class of wage earners whose members are fairly homogeneous in quality; that is to say, a group so alike in function and capabilities that one man can be substituted for another without material loss to the industry. In the second place we simplify our problem by assuming at the start a condition of free competition and a perfect labor market and developing the theoretical principles which would operate under such confessedly artificial conditions. Under these assumptions, there will tend to be a uniform wage rate for the group, determined by demand and supply. The qualifications required to adjust this simplified picture to the complex facts of the industrial world can be made after we have analyzed the forces which tend toward a uniform wage rate within this section of the labor market.

The demand for labor. Let us first inquire into the forces which control the demand for labor. Wages are paid in the first instance by profit-seeking employers, who voluntarily undertake to give the workers a money income in exchange for their services. Accordingly a study of the demand for labor leads us to inquire what determines the demand of a given employer and the price he is willing to pay. To the employer wages are costs of production similar to the costs involved in equipping his enterprise with capital instruments. He assumes the cost of paying wages for the same reason which impels him to burden his future income with

interest charges when borrowing capital. That is to say, he expects the workers to create the income out of which wages will be paid, just as, when borrowing capital, he expects to increase his income by an amount at least equal to the interest charge. In both cases he has but one source from which to recover his costs; that is, through the sale of the product of capital and labor. The employer's demand for labor is really a demand for the product of the wage-earner; the money value which he can place on labor is derived from the price which he must accept when selling the output of his enterprise to consumers on the market. We may view the employer therefore as an intermediary who stands between the worker and the market, assuming the task of selling what the worker has produced and transmitting to him the proceeds of the sale.

But recognition of the fact that the demand for labor is based on the laborer's productivity is only the first step toward explaining the forces which determine the wage rate. It would appear that the workers might demand the entire income of an enterprise as their just share, since no portion of the product could have been produced without their assistance. That this does not happen is due to the fact that all employers are not obliged to choose between the alternatives of hiring the whole complement of their labor forces as a single unit or of hiring none at all. Under conditions of competition such as we are considering, workers are employed as individuals, just as land and capital are obtained and employed in successive increments. Now, although a simultaneous withdrawal of all the workers from industry would certainly reduce the product to nothing, no such effect will result from the withdrawal of one man's labor or from the withdrawal of the labor of that number of workers which constitutes a normal labor unit. The specific productivity of a single member of a homogeneous group is equivalent to the difference in the income of the enterprise which is attributable to the efforts of this individual. It is the desire to obtain this increment of income which creates the employer's demand for units of labor and sets the limit to the amount he is able to pay for the laborer's service.

Diminishing productivity as applied to labor. When we examine the productivity of the wage earner a little more closely, we find that it conforms to a law of diminishing productivity. The contribution which each unit of labor makes to the income of a business enterprise tends to fall as the employer expands his labor force, other factors remaining constant. It is true that the principle of diminishing productivity may not become apparent during the early stages in the expansion of the labor supply. More efficient division of labor and a better combination of the different sorts of service which comprise the labor factor in an industry may even result, for a time, in an increase in the average workman's contribution to the corporate income. But this cannot go on indefinitely. If a modern factory with a given capital equipment increases its total complement of workers — executive, clerical, manual — a point will be reached in the process of expansion when additional units of the labor force will be less effective in increasing the enterpriser's money receipts than were the preceding units.

There are two reasons why this is so. As was shown in our study of interest when the quantity of capital increases while the coöperating factors remain fixed in quantity, the physical product of the combination tends to fall as measured in proportion to the number of units of the increasing factor. So it is with labor. The physical product per unit of labor will eventually decline as the number of laborers is increased. In the second place, beyond certain limits increasing quantities of output must be sold at lower prices. This double reason makes it impossible for any employer to go on expanding his labor force — land and capital constant — without eventually noting a decline in the amount of income which he can attribute to one unit of labor. What we have said with regard to the entire labor force is equally true if applied to one occupational group within the industry. If the increase takes place within this one group — let us say, the machine operators — while all other factors remain constant, there will be evidence of the same principle of diminishing marginal effectiveness as applied to the increasing group.

Law governing an employer's demand for labor. We now have before us the considerations which govern the demand of the employer for the labor of a particular type of worker and for each of the different classes which comprise the labor supply. In the first place, the hiring of workmen being optional with the employer, he will offer employment only on such terms as are remunerative to himself. Secondly, since he can recover what he expends in wages only through the sale of his product, his appraisal of the economic worth of a workman is really a reflection of the prices fixed upon his product by consumers' demand. His test of the productivity of the worker is the contribution of that worker to the money income of his business; but this whole income, together with the fraction attributable to the workman, is controlled fundamentally by the market price of his good. Finally, as he increases the amount of any one type of labor or of all types together, the marginal man in each group will contribute less and less to the corporate income. With these facts in mind, it is not difficult to state in abstract terms the law which determines the employer's demand for labor. Assuming, as he does, a fixed rate of wages, *the employer will add to his labor force until the contribution of the marginal man in each group to the income of the enterprise is no greater than his wages.* In common sense terms this means that the employer will continue to hire men as long as he gains by so doing, and he gains as long as the value of the marginal man's product is in excess of the wage rate. Evidently the lower the wage rate, the greater the number of laborers that can be employed before the limit is reached; from which it follows that *the number of laborers that the employer will hire at any given time varies inversely with the rate of wages.*

Wages are advance payments. One important technical feature remains to be considered. The workers will want their pay in advance of their employer's sale of their product. Generally at the end of each week the laborers will ask to be paid for the services performed up to that time. But in the majority of cases the products upon which they have been working are not in a completed form and ready for the market at the time the wages are

paid. This phase of the matter was touched in our study of interest, where the capitalistic process was described as a succession of advances to workers. If wages are to be paid at short intervals, some one must wait, through a period more or less protracted according to the nature of the product, before the money advanced to the workers can be recovered from the sale of the product. In the modern organization this function of waiting has been delegated to the employer. But the employer will not thus make advances to laborers unless he is paid for his waiting; that is, unless he receives interest upon the advance payments. The factor which determines the wages he can pay is therefore not the full value of the product at the future date when realized, but the present worth of that value at the time when the wages must be paid. The wages which the employer can afford to pay the marginal worker therefore will fall short of the value of his contribution to the product of the business by the amount of the interest charge.

The general demand for labor. Thus far we have been viewing the problem of wages through the eyes of the individual employer, to whom the wage rate appears as a price more or less beyond his influence. Broadening our view to take account of the whole range of industrial enterprise as causing a demand for the labor of propertyless men and women, we see the existing body of laborers seeking employment over this great field of industrial activity. On the whole, they are successful in their search for jobs; for involuntary unemployment is the exception and not the rule in modern society. When the members of any labor group apply for employment to one industry after another, they soon learn that the conditions which govern the demand of the individual employer hold good universally. Because of diminishing productivity, neither a single employer nor all employers combined can provide jobs for increasing numbers of this labor group except at successively lower rates of money wages, so long as the supplies of coöperating factors — land, capital, other types of labor — remain unchanged. *The rate of money wages for any group of workers will tend to equal the discounted value of the product of the marginal man in that group.* This we may call the law of the demand for labor. It corresponds

closely with the general principle that the price of anything, in a competitive market, tends to be equal to its marginal utility to the purchasers.[1] So wages, the price of labor, tend to be equal to the marginal utility of labor to the employer; that is, the discounted value of the product of the marginal unit.

The supply of labor. But the foregoing is only half of the law of wages, since the marginal product of labor is itself a function of the number of laborers. Our problem is not solved until we discover the law of the supply of labor. This, however, in the light of what we know of supply in general, is a comparatively simple task. In general, other things being equal, the quantity of any good which will be offered for sale varies directly with the price. So the number of laborers that may be secured in any particular labor group is dependent upon the rate of wages offered. While labor thus conforms to the general law of supply, it should be noted that the circumstances back of the supply of labor are in many respects peculiar. The amount of labor in general available at any given time is a comparatively fixed quantity, not susceptible to great variation in relation to the rate of wages. In other words, the supply of labor is comparatively inelastic. Even if we consider a particular homogeneous labor group, the inelasticity of supply is only slightly modified, since, as will later appear, laborers are not easily drawn from one group into another, even by the inducement of considerably higher wages. Moreover changes in the number of laborers take place slowly, and there is only in a remote sense such a thing as a cost of production of labor. Leaving for later consideration questions such as these, we may reach our present goal by accepting the general law that the supply of labor is similar to the supply of things in general in that the quantity offered varies directly with the price, though in the case of labor the variation is only moderate.

Demand and supply; the market rate of wages. We are thus in a position to represent our conclusions graphically by means of the familiar curves of demand and supply, as in Figure 6. Quantities of labor of a given type are measured on the axis ox; wage rates

[1] See Chapter XVI.

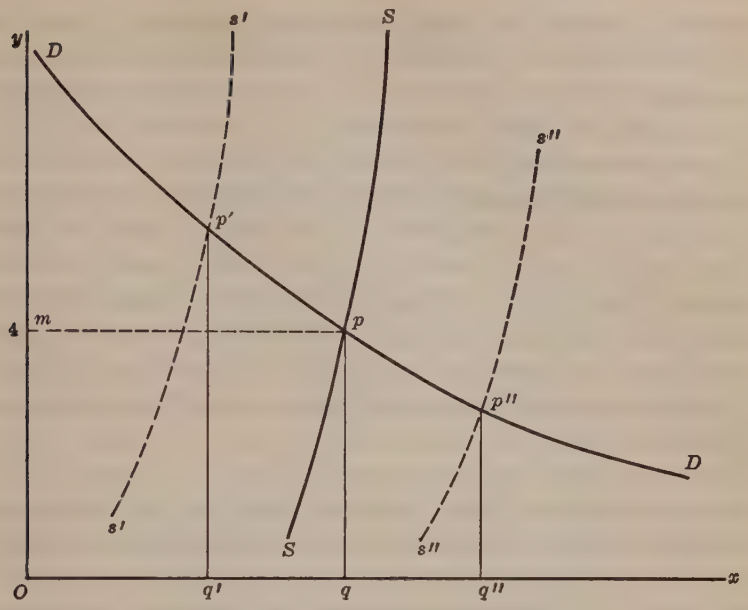

FIG. 6.

paid for different quantities of this labor, on the axis *oy*. The demand curve *DD* represents the relation between the rate of wages and the quantity of labor that would be employed. The height of any point of the curve, as indicated on the axis *oy*, shows the maximum rate of wages which would be paid for the quantity of labor as indicated by the abscissa of the point. The distance of any point to the right of the origin, as measured on the axis *ox*, shows the maximum quantity of labor that would be employed at the price indicated by the ordinate of the point. Since the wage rates offered by all employers for different numbers of workers or the number of workers that would be hired at different wage rates are determined by the discounted money value of the marginal products of these workers, this curve may also be called the marginal productivity curve of labor. To illustrate, given the quantity of labor *oq*, the line *pq* represents both the wage rate offered for this number of laborers and the value of their marginal product. Or, given the wage rate *pq*, the line *oq* represents the

number of laborers that would be hired at that wage. The law of diminishing marginal productivity of labor is represented by the downward trend of the curve. The curve SS is the supply curve, its steepness indicating the inelasticity of the sale of labor.

It is apparent that the market rate of wages depends upon both the demand and the supply of labor, being determined by the point of intersection of the two curves. The rate of wages will vary directly with the productivity of labor and inversely with the number of the laborers. We arrive at the following general law of wages. *The rate of money wages for any group of workers tends to be determined by demand and supply of labor at that point at which it is equal to the discounted value of the product of the marginal man in that group.*

Influence of the supply of labor. If labor of a certain kind — say common factory labor — is scarce, each unit of it will play an important part in industry; its marginal productivity will be high, determined by the intersection of the supply curve well to the left, as at p'. Employers competing for its use will offer high rates of money wages. If the workers are able to take advantage of this competition, they will so distribute themselves among the different enterprises that all employers pay the same rate of wages.

On the other hand let us assume that, after the wage rate has settled at some point — say four dollars per day (pq on the graph) — an influx of immigrants causes a sudden increase in the number of workers qualified to perform the duties of factory labor, pushing the supply curve over to $s''s''$, conditions of demand remaining unchanged. In order to avoid unemployment, these new workers must offer their labor at lower rates of wages, since their entrance into industry will have reduced the marginal productivity of their group. Accordingly we may picture the process of competition working out a new distribution of labor until the point is reached again where all are employed, where the marginal product is similar in all enterprises; and where a common wage rate ($p''q''$) measures the discounted value of the marginal man's services to each employer.

Consider now the effect of this change upon the welfare of the

older group of workers. Presumably they are working now as hard and faithfully as before, when the wage rate was higher. On ethical or moral grounds, they cannot see that they are deserving of less pay; yet they are unable to resist a wage cut. In any group of workers so homogeneous that one man may be substituted for another, each member is the marginal laborer. If an enterpriser employs fifty such men, the importance of any one of them to him can be measured by the difference in his income caused by having fifty at work instead of forty-nine. The fiftieth man may be identified as any one of the group, for the absence of any one would merely have the effect of reducing the working force to forty-nine. Similarly if a new man is added to the group, increasing it to fifty-one, each of the members of the original force will then occupy the same position as the fifty-first. Since the marginal productivity will fall with the influx of new workers, each of the former employees will lose a portion of his importance as a producer and, though he expends as much energy as before, will be worth less. It is often said, when speaking of such a state of affairs, that the employer in cutting the wage rate takes advantage of the competitive weakness of his employees to impoverish them through no fault of their own. But we see that the decline of the wage rate does not occur because the employer has changed his wage policy; he continues as before to pay each of his workers what he is worth as a producer — the discounted value of his product. The economic law of wages does not attempt to appraise the moral worth of the worker or to give him an income commensurate with his needs as a human being. It guarantees no man a " decent living wage " but measures his specific importance as a contributor to the social income.

Relation of the individual employer to the labor market. Returning now to the position of the individual employer in relation to the labor market, let us assume again that the wage rate has settled at the point indicated on the graph — four dollars per day for an amount of labor indicated by the quantity *oq*. In this situation we will find the individual employer occupying that position in the labor market which we pictured earlier in our dis-

cussion. He will accept the rate as a necessary cost of production and will concern himself with the problem of deciding how many workers to employ at that rate. It will be to his advantage to expand his labor force until the discounted value of the marginal man's product is four dollars. Up to this point each worker will contribute income in excess of his costs. For the entire labor market this fact is indicated on our graph by the area *mpD*. This part of the income is available for distribution to the other productive factors in the industry, including a reward for the managerial and risk-bearing functions of the enterprisers. Reduced to small scale, this surplus income from labor in excess of costs appears in each individual enterprise and provides an inducement to each employer to expand his labor force until its discounted marginal product is no higher than the wage rate. Unconsciously the individual employer will aid in fixing the market rate of wages at this point. For if he were to discontinue his enterprise and withdraw from the market, the workers whom he releases must seek employment elsewhere, thereby (in theory) lowering the marginal productivity in other enterprises and reducing the wage rate. On the other hand, if this employer is enabled by some development in his business to increase his labor force without reducing their marginal worth, his increased demand for labor could be satisfied only by diverting wage earners from other industries, thus raising marginal productivity in general and increasing the wage rate.

In strict theory no employer can obtain so large a supply of any type of labor that its marginal productivity in his enterprise falls below its marginal productivity in the enterprises of competing employers. If this should happen, a second employer would be able to offer higher money wages than the first and would draw labor away from him. When competition on both sides of the labor market has worked out a stable distribution of the labor supply, some employers will have more of this type of labor, some less; but, in theory, the last man in the working force of each employer will contribute substantially the same amount to the income of his enterprise.

Practical qualifications. Dependence of different types of labor. This is the theoretical view of the matter. Between the theoretical statement and the actual result in practice there will be the discrepancies which must always occur when theoretical analysis is unable to take full account of all the circumstances of any practical case. This matter has been called to the reader's attention in previous chapters. In the case in hand our law of general money wages is obviously an abstract statement of a tendency rather than an explanation of all the diversities which may be observed in the business world. We must now add qualifications to the law in order to take account of certain facts of experience which, if unexplained, might throw doubt upon the validity of our conclusions. In the first place it never works out that the different enterprisers enlarge their several labor forces to the point where marginal productivities are precisely equal. The number of workers of a certain type which can be employed profitably in a given industry is affected, not only by the marginal productivity of these workers, but also by the necessity of maintaining an effective proportion between the different types of labor in the enterprise. Complex division of labor creates a situation in which no single type of labor can be employed without necessitating the employment of other types. For example, an industry employing one bookkeeper might not have need for the services of another at any rate of wages, so long as its present scale of operations remained unchanged. To make it possible for a second bookkeeper to produce anything might require the employment of a dozen machinists, a foreman, a shipping clerk, and two or three other types of labor, none of which could be employed profitably at their prevailing wage rates. In a certain large factory producing tennis shoes, one inspector is employed in conjunction with twenty-five workers in other lines; one cutter works with half a dozen stitchers, etc.

These examples illustrate the fact that the demand of any employer for a given type of labor is limited by his force of coöperating types, which are limited in turn by the wage rates of these other labor groups. He tends to increase the number of workers in each group up to the point where their marginal productivity equates

with the wage rate, but he may stop short of this point because of his inability profitably to expand the scale of his industry. Obviously the small business is much more affected by this consideration than is the large one which, with adequate supplies of capital, can take full advantage of the gains inherent in complex division of labor. When however we fix attention on a single occupational group, it is probably true that in no individual enterprise could we find that precise equivalence of marginal productivity with the wage rate which our law of wages implies.

The sphere of bargaining. As we have stated the law, it would appear that the rate of wages tends to settle in some automatic and frictionless fashion at a point which measures the marginal productivity of the workers. This is because we have thus far assumed conditions of perfect competition. But in practice the rate for any group of wage earners in any given industry is settled by bargain between employer and worker and fluctuates according to the relative bargaining power of these two parties to the wage contract. The actual rate in a chosen industry or for a chosen section of the laboring population is very likely to settle at a lower point than that determined by economic law and to remain at this point for periods of considerable length; on the other hand, it may rise above this point for short-run periods in isolated enterprises or localities. We cannot undertake to follow out in detail these complex variations, but it will be profitable to ascertain as nearly as we can the limits of this sphere of bargaining within which actual wage rates fluctuate.

Wages above marginal product. Circumstances may arise in which the employer will voluntarily pay some of his workmen more than they are worth, thus leaving himself worse off than if he had refused to hire them at all. Employers are not devoid of feelings of compassion and do at times mix charity with business. Furthermore it may be to the positive advantage of a given employer to accept a loss on the wage contract for a short time in order to avoid a greater loss involved in dismissing his employees. During periods of depression the enterpriser will look forward to a recovery of the market which will set his industry in full operation again.

By keeping an experienced labor force together he will avoid the considerable loss involved in hiring new workers, breaking them in to the routine of the plant, and training them in his methods of management. It may be worth his while to do this even at the expense of paying his workers for doing nothing. These are but indications of the motives which may impel employers voluntarily to pay wages above the rate set by economic law.

It is also possible for the workers, under certain conditions, to force an employer against his will to pay them more than their marginal worth. By threat of strike a group of workers which holds a temporary monopoly on the supply of labor may compel the employer to make a choice between the alternatives of paying his labor force more than it is worth or of suspending operations. If an employer could always close his plant without loss, it is not probable that he would ever accede to such demands. But this he cannot do. To suspend operations will at least wipe out his own wage income — the reward of his labor of management. If compelled to do so, he may consent to raise wages above the marginal worth of the workers, surrendering to them a part of his own wage income up to the point where his share in the revenues of the enterprise has disappeared. Even here we do not have a final limit to the advance of wages, for it is conceivable that aggressive bargaining by the workers may raise the rate so high that the enterpriser is not only not receiving payment for his own labor but is actually operating at a loss. Most employers have fixed costs which will run on unchanged even if they close down their industries. It is the lesser of two evils to continue running at a current loss, provided the loss is not as large as the total of the current fixed charges. This condition is of course self-limiting; its ultimate limit will be reached when the employer's equity in the business has been wiped out. But it is quite possible that an employer, operating under heavy fixed charges and with a large investment of his own funds in his enterprise, can be made to bear a loss of this kind; and there are enough cases of employers who have been forced into bankruptcy as the result of a series of losses on the wage contract to make it clear that workers holding a

strategic bargaining position may turn this fact to their advantage.

We see then that there are three resistance points to a rise of wages. The first and most important is set by the law of marginal productivity. It is rarely " good business " to pay a higher rate than this, and no employer will voluntarily do so unless it be as a temporary expedient in abnormal circumstances. The second check on rising wages is exerted at the point where the employer's profit disappears. This is the maximum limit to the range of bargaining except in industries materially affected by fixed costs. The highest conceivable rate under any conditions is set at the point which causes a current operating loss equal to the loss involved in closing down the industry. Wages will rise above the first of these points — that determined by marginal productivity — only when the workers are possessed of some unusual bargaining advantage. Even then such a rise of wages can appear only in isolated cases and for short run periods. When his wage rate rises to the second of our resistance points, the employer will soon give up the struggle. No employer will be content to work indefinitely without reward. The duration of a wage rate at a level between the second and third resistance points is automatically self-limiting by inevitable insolvency. We may safely conclude that advances of the wage rate above the point of the discounted marginal product will appear only as exceptional and short-lived occurrences.

Wages below marginal product. Having discussed the possibility of a wage rate higher than the discounted marginal productivity of the workers, we may now consider the reverse condition. There are reasons why we cannot conclude that wages cannot normally fall below the point determined by this law. Complex division of labor has resulted in the splitting up of processes into minute fractions of the whole, making it in many cases quite impossible to fix a precise value on that fragment of the composite product which is attributable to the labor of a single workman. Accordingly there is no clearly defined test of the marginal value of the product of any group which all parties to the wage contract will accept as fair. Since this is true in individual cases, may we assume

that these valuations result in some way from the operation of competitive forces in the general industrial world and that each employer adopting this rate will pay to the worker his maximum value to industry as a whole, even if this does not conform to his exact economic worth in the employer's own enterprise?

Such an assumption would rest upon the belief that competition works as freely on one side of the wage contract as on the other. But this is untrue in actual life. For wage earners, as one will be convinced by even a superficial examination of the facts, are not usually in position to take full advantage of competition among employers. To do so they must be able to play off one employer against his competitors, to move about at will from one locality to another, to discard service in one industry and undertake employment in a different one whenever it is to their immediate economic interest to do so. All of this requires a degree of mobility in the ranks of labor which the average workman does not possess and which he is prevented from acquiring by (to mention but one factor among many) the mere cost of transportation for himself, his family, and his household effects. It must be remembered that the labor market requires the workman to carry his labor to the employer; it does not require the employer to bring the job to the workman. Perfect competition in the labor market then rests on physical mobility among laborers, to attain which they must overcome obstacles of financial expense, of home and community ties, of custom, tradition, and sentiment.

There seems to be no doubt that the bargaining weakness of the workers has often resulted in a wage rate lower than their marginal economic worth. Indeed in some cases it is probable that groups of laborers have been underpaid in this sense through long periods, if not through the whole of their productive lives. The minimum wage laws of certain of our states caused the wages of women to rise by as much as eighty per cent without affecting the prices of their products or causing unemployment. That such a change as this could have occurred without being followed by one or both of these consequences can argue only that these workers were receiving less than their economic worth prior to the passing of the laws.

Similar incidents have occurred in trade union history, and they point to the same conclusion. The decline of the wage rate below this theoretical point therefore is not necessarily an abnormal and temporary occurrence, as is the rise of wages above it, but may be a natural result of the unequal bargaining power of the two parties to the wage contract.

It would aid us to limit the range through which the wage may fall if we could ascertain with precision lower limits to the sphere of bargaining comparable to the resistance points which check a rise of wages. In a general way, two such resistance points can be marked out. The first of these is determined by the wage-earners' traditional standard of living. Among the skilled labor groups especially, self-respecting workers will reject offers of employment at rates which spell a degradation of their standard of life and a loss of social status. Such behavior is not to be found among all the workers, nor in any case will it suffice to resist permanently a fall of wages. If employers are in a position of bargaining advantage and determined to press their power to the utmost, they will eventually break through any resistance founded on a traditional mode of life. A second and ultimate lower limit is reached when the rate falls to the minimum of subsistence. If starvation is to be their fate in any case, free men will prefer to starve at leisure rather than while at labor to another man's profit. Even if they were content to work for less than subsistance wages, the duration of such a wage rate would be limited by its inability to maintain the supply of labor.

Lack of uniformity in wage rates. For purposes of exposition, we have thus far assumed that the workers in each specific line of occupation are so homogeneous that one can be substituted for another at will. This assumption has led to the conclusion that all members of a given occupational group will receive the same rate of wages; that all carpenters will be paid alike, all stenographers, all machinists, all school teachers, etc. Everyone knows that this is not strictly true in fact. The wage rate varies for different members within an occupational group and also for different localities.

Considering now the lack of uniformity in the earnings of different members of the same occupational group, it is obvious that this fact may sometimes be due to relative difference in bargaining power such as we have discussed above. Women are habitually paid less for the same service than men. Negro labor in the United States is similarly underpaid in comparison with white labor, and the same is true of recent immigrants in comparison with workers who have acquired our culture and standard of living. The timidity of women workers, their inability to take advantage of potential competition among their employers, their indifference to opportunities to improve their immediate economic status weaken their bargaining power. In the case of colored labor social prejudices and the improvidence of the negro, in the case of recent immigrants their ignorance and contentment with a traditionally low standard of life, go far to explain the fact that these groups do not receive the same pay for a given service as other workers. These factors influence wage rates indirectly through their effect upon the bargaining power of the workers.

But of chief importance in explaining lack of uniformity in the earnings of different members of the same labor group is the fact that no labor group is quite homogeneous in productivity. If we select any occupation at random, we will observe that all its members do not have the same productive efficiency. Noticeable differences appear in the number of tons of coal credited to the account of different miners at the end of an eight-hour day. The tenders of machines in a cotton mill will differ somewhat in the speed with which they perform their operations and in their ability to sustain a given rate of production through the hours of the working day. Other examples could be cited. Closer observation will show that these differences in productivity are greater in the case of some labor groups than others. In the crudest forms of labor the difference in the output of one member of the working force as compared with another is usually very slight. So also with the great army of workers in the factories; there is a marked uniformity in productivity among the millions of machine operatives in the great industries. The differences are greater among

skilled craftsmen, where native ability and training play a relatively large part in the mastering of the trade. And they are greatest of all in the highest labor groups, the learned professions, and the superior grades of managers and superintendents.

Forces tending toward uniformity: machinery. Certain tendencies in the modern industrial world are working toward greater homogeneity of large labor groups and greater uniformity in the wage rate applying to these groups; certain other forces tend in the opposite direction. Among the former the one of greatest importance is the incessant march of machine technique. Since the Industrial Revolution, the passing years have witnessed a gradual but inexorable encroachment by the machine on the domain of the skilled craftsman, and this progress is still in motion. As was pointed out in an early chapter of this book, complex division of labor has proved favorable to continuous invention and to the substitution of mechanical processes for the human hand. One craft after another has been broken up into fragments, and a variety of power-driven capital instruments has displaced the versatile artisan who formerly performed the entire process. As the skilled trades are conquered by the machine, the workers tend to fall into a great class of machine tenders, performing simple, routine operations which give but narrow scope to differences of ability and training. The rate of production as well is taken out of their control, being regulated by the flow of power which drives the machines. This increasing mechanization of industry tends to make the workers mere interchangeable parts in the gigantic productive mechanism of society, levelling down their differences as individuals and stamping them with a common mould.

Trade unionism. Of similar effect is the growth of labor organization. Trade unionism flourishes especially among the remaining branches of the skilled trades, those groups of workmen which would disclose large individual differences in productivity if governed by the free play of self-interest. The essence of trade unionism is conscious and intentional coöperation for purposes of common benefit among the members of these groups. One standard device of the union is *collective bargaining* — the demand that

employers hire the entire working force as a unit according to terms set in a single wage contract. Another standard device is the uniform rate, or the *union scale* of wages. Each employer who deals with the union is obliged to pay a standard rate to all members of the craft and to accord to each the same treatment in respect of hours of work. As the trade union grows in strength, this standardization of wages spreads from one industry to another until it covers the entire nation. Though not officially prohibited by the union, the practice of paying rates of wages higher than the union scale to selected individuals is frowned upon, and the public opinion of the group tends to prevent any member from accepting favors of this kind. Wherever the nature of the occupation lends itself to such practices, the unions tend to standardize the productivity of their members as well as their rate of wages. If their output is measured by units of product — *e.g.* the laying of so many brick per day, the construction of so many moulds for iron castings, the turning of so many pieces on the lathe — organized labor will set up a certain number of these units as constituting a day's work. If the product is not subject to such enumeration, the same result is reached by setting a " decent " pace of work which becomes the norm for all good union men. The motives which lie behind these practices are complex, consisting chiefly of a desire to stifle individual competition, which might be utilized by employers to disrupt the union, and to defend the workmen against the menace of unemployment. Without attempting to discuss the problem in all its bearings now, we may observe that the effect of such practices as these is, like the effect of machine production, to promote a tendency toward uniformity in productivity and earnings for large groups of workers over wide territorial areas.

Systems of wage payment. Where differences of individual productivity persist despite these tendencies toward uniformity, the business world attempts to take account of them by adjusting the method of wage payment to serve the double purpose of increasing the reward of the more productive workmen and stimulating the individual's desire to increase his productivity. The perfecting of these devices of wage payment and labor management

has led in recent years to the development of a new industrial science and a new profession for business executives. The traditional method of wage payment has been to base the rate of wages on a unit of time — the hour, the day, or some longer period. This method is difficult to adjust to individual differences in productivity and tends moreover to stifle the ambition of the better workmen. Employers who use the time base for wage payment generally attempt by disciplinary methods to induce from their workmen a rate of productivity at least as high as the wage rate and as much more as they can get. Slight differences in productivity are disregarded, and a " drive " policy of labor management is relied upon to keep the slackers up to the mark. In cases where differences in individual productivity are so marked that they cannot be overlooked allowances are made for the superior efficiency of the better workers by means of differentials, which measure roughly their specific productivity as compared with the average man. These allowances do not always take the form of increases of money wages. The faithful and industrious members of the labor force may be given privileges in the form of the more attractive jobs, or they may be the first to be employed and the last to be discharged in periods of business fluctuation.

Piece rate wages. The difficulty of adjusting the time wage to differences in individual productivity and the tendency of this system to impair the workers' efficiency have caused modern business executives to devise new systems of wage payment which will be free from these disadvantages. These systems vary in detail, but they all embody a common principle of *piece rate wages*. The first step in establishing a piece rate system is to resolve the product of the worker into a definite, measurable unit called the " job." The unit of measurement must of course be adjusted to the conditions of the particular enterprise. An iron foundry may adopt as a " job " a standard mould or casting; a coal mine, a ton of coal; cotton mills may use for this purpose a yard of their standardized product; the shipping department may have for one of its labor groups the construction of a box of standard size; for operators of automatic or power-driven machines, the " job "

may be a certain number of units of a processed product. The variations in detail are almost infinite, though the principle is the same in all cases. The next step, after settling on the " job," is to discover by experiment how many of these units the average workman can produce in a given length of time. This is usually done by selecting a subject for the experiment and timing him with a stop watch. With this information at hand, the final step is to fix a rate or price on the " job " which will return the average workman a rate of money wages equivalent to the customary earnings of men of equal skill in the general world of industry.

Under the piece rate system the workmen are invited to make the most of their opportunities, to produce as many units as they can in the course of a working day, and hence to prosper in accordance with their various capacities. Obviously this system of wage payment tends to break up the homogeneity of a labor group and to accentuate the difference in individual productivity and in the money earnings of workmen performing the same sort of labor. Refinements of piece rates, commonly known as *task and bonus systems*, modify somewhat these tendencies to increase individual variations in productivity. These improved systems of wage payment also vary in detail, but many are alike in that they guarantee a standard rate of *hourly* earnings to all the workers, while rewarding those whose output exceeds the normal daily product by paying bonuses proportioned to the amount of the excess output.[1]

Piece rates and the economic law of wages. These developments in the science of labor management do not upset the operation of the economic law of wages. In the first place, the piece rate system is applicable only over a relatively small range of industrial activity. It cannot be applied at all to workmen whose product is not tangible and definitely mensurable, and this is the case with by far the major part of hired laborers. Moreover, even in those industries which have adopted the piece rate system, the

[1] For more detailed exposition of the piece rate and task and bonus systems of wage payment see C. B. Going, *Principles of Industrial Engineering*, McGraw-Hill, 1911.

rate of wages is not divorced from the principle of marginal productivity, but is more closely assimilated to that principle. The basic piece rate is frankly set to harmonize with the customary earnings of the group in question and is therefore governed by the general experience of the business world where marginal productivity is the norm of money wages. Though it results in differences within the shop which employs it, these differences are almost automatically adjusted to variations in individual productivity. The working force, instead of being merged into a common group with a single marginal productivity and a uniform wage rate, is broken up into smaller groups, each with its own peculiar productivity and its own rate of wages. The earnings of the general run of factory operatives will tend to conform to the general wage rate of their class, while the superior workmen will receive allowances which measure the excess of their products when compared with that of the average man.

Variations in local wage rates. Wage rates not only differ for different members of a given occupational group in the same place, but they vary between localities. To some extent these differences may be attributed to the effect of the bargaining factor which we have discussed. For various reasons the bargain which fixes the actual wage rate within the limits described on a preceding page terminates variously in different localities. Enterprisers are not alike in their efficiency as managers, and consequently different business concerns vary in profitableness. Some units of a given branch of industry are equipped with better machinery or possessed of more favorable markets or otherwise favored as compared with competing units. These differences affect the ability of an enterprise to pay wages, and they also affect the aggressiveness with which the employer bargains for a saving of cost in the wage contract. Employers with the smallest profit and those operating without profit or at a loss are proverbially the most aggressive bargainers. On the other side of the labor market, differences in the character of the working people, in their standards of life, in the degree to which they are unionized, and in other local conditions materially affect the bargaining strength of different groups

of laborers. Such differences in local conditions modify the resistance points in the fluctuation of wages and cause local variations from the norm. The effect of the bargaining factor upon local wage scales is much more noticeable in the case of some types of labor than in that of other types. The wages of women in any given locality are always much affected by their lack of bargaining aggressiveness; hence these local wage rates are apt to vary markedly as between different business centres.

For example, the wage rates for one group of women workers in hosiery mills (loopers) showed the following local variations in 1922; Connecticut, 34 cents per hour; Georgia, 24 cents; Indiana, 39 cents; Massachusetts, 37 cents; New York, 30 cents; North Carolina, 16 cents. The spread between the lowest and the highest of these rates is almost 150 per cent. Other forms of female labor showed similarly wide local variations, as did also the wages of negro workers, of unskilled immigrant labor, and of other labor groups similarly affected by their bargaining weakness. On the other hand, the wages of skilled male occupations disclose much smaller local variations, and those of unionized male workers are still more uniform. In 1923 the wage rates for unionized carpenters were standardized at $1.00 per hour in such widely separated cities as Chicago, Cincinnati, Grand Rapids, San Francisco, and Shreveport, and they varied for all cities of considerable size by an amount not exceeding 30 per cent in extreme cases. At the same time the wage rate for non-unionized carpenters was as low as 50 cents in Savannah and as high as $1.00 in Tucson. These differences cannot be attributed entirely to the effect of bargaining, but there can be no doubt that this factor goes far to explain the extreme local differences in the case of workers who are notoriously weak bargainers.

The basic cause of local variations in wage rates however is the fact that the labor market is broken up into small fragments between which the forces of competition either do not operate at all or operate haltingly and with much friction. If the student will hark back to our study of interregional trade in an earlier chapter, he will recall that for commodities in general the prices in different

places cannot differ by more than the cost of transportation between the places in question and he will further recall that labor was there referred to as an exception to this rule on account of its extreme immobility. If all the workers in a given occupation could and would move at will over the whole industrial field, local wage rates would tend toward much greater uniformity. But in reality each business centre forms a labor market of its own where the relation of supply to demand may be quite different from that of other centres and where consequently the marginal productivity of the workers will also be different. Local wage rates tend to vary in rough proportion to the cost of living; they are lower in small towns than in the metropolis, in Southern than in Northern cities, in the Middle West than on the seaboard.

This fact lends support to an erroneous theory that the wage rate is determined by living costs. But local differences in the cost of living can affect wages only through their influence on the distribution of the labor supply and on the bargaining power of the workers. The cheap living which they enjoy in a country town or in a small manufacturing city not in close communication with the general business world may be an element in the unwillingness of a labor group to move to the metropolis where wage rates are higher. In this case the low wage rate will measure a low marginal productivity. Low living costs may also perpetuate a traditional wage scale in a given locality which is probably lower than the economic worth of the workers in that place. Contentment of the workers with a customary mode of life weakens their aggressiveness as bargainers and, in conjunction with their unwillingness or inability to move elsewhere, operates to allow the local wage rate to fall to a low point.

In concluding this topic it should be observed that modern developments are tending to reduce these local variations in wage scales by broadening the labor market and removing differences in the bargaining power of different labor groups. Cheap transportation increases the physical mobility of the workers; newspapers and magazines with nationwide circulation stimulate similar desires in widely separated labor groups and stir the ambition of

the more backward; education breaks down the barriers of provincialism and traditional standards of work and pay. Negro labor is beginning to migrate to the north; the small towns and back country cities witness an annual exodus of their younger laborers to the chief business centres; immigrants quickly assimilate our standard of living, and their children appear to be born with the American's voracious demand on life and his restlessness to better his condition. The spread of unionism promotes the mobility of local groups and at the same time stiffens their bargaining power. The peculiar weakness of female workers is disappearing as it becomes an established custom for women to seek a life of economic independence and a permanent place in the industrial world. These forces and many others of similar nature broaden the area within which a common wage rate holds sway.

Marginal productivity of fundamental importance. Because of the complexity of the labor market it has been necessary for us to introduce various qualifications into our law of wages in order to bring it into rough conformity with the facts of industry. We have recognized the fact that the actual wage rate is determined by bargaining and have marked out the limits of the sphere within which the incidence of the bargain will fall. We have taken account of the existing lack of uniformity in the wage rate for members of a given occupational group and have accounted for this by referring to differences in bargaining power and especially to differences in individual productivity. We have also noted local variations in wage rates for a given type of labor and have shown that these result from local differences in relative bargaining power and especially from maladjustment in the labor supply.

But this discussion of complicating factors must not cause us to lose sight of the essential truth embodied in the economic law of wages. The wage bargain revolves around the point of discounted marginal productivity as a norm; actual wage rates cannot long persist above this point, and all the developmental tendencies in our society work against their falling far or remaining for long below this point. Homogeneity of labor groups is promoted by various tendencies which we have noted; other forces are broaden-

ing the scope of the labor market and promoting a more even adjustment of the local supply of labor to the demand for it. We may conclude that no device of wage payment and no alterations in the method and procedure of the labor market can materially affect the fact that propertyless men in a society sanctioning private property and individual enterprise will normally receive wages approximating the value of their specific contributions to society's income. If they are members of groups whose individuals have much the same rate of productivity, the value of each man's contribution will be equated, through the principle of substitution, to that of the marginal man in the group. If they must receive their share of the social income before the sale of their product, their money incomes will be reduced by the discount process through which present and future values are equated. Upon these valid considerations rests the economic law of general money wages — the law of the discounted marginal product of labor.

Other types of labor income. Our exposition of the law of wages was made with sole reference to the hired worker; but our conclusions will serve to explain the money incomes of those who are their own employers. Here the typical case is complicated by the fact that the independent worker must almost of necessity be an owner of capital and to some extent a manager of enterprise. The physician, the lawyer, the barber — examples could be multiplied indefinitely — possess productive capital of one sort or another and often land as well; their money incomes are a mixture of rent, interest, and wages, no single element in the income being clearly defined in contrast with the rest. It is not impossible, however, to arrive at a rough measurement of that part of their income which is attributable to labor, by deducting allowances for capital and land at the market rates of interest and rent.

In the case of independent workers, no intermediary stands between them and the market. Their money incomes are drawn directly from the buyers of their products. They rise with a rise in the price of their product and fall as prices fall, tending to settle at any given time at points which measure the marginal utilities of their services to consumers. The fact that these groups are

less homogeneous than are groups of hired workers, that there are for example relatively large differences in the productive efficiency of physicians, attorneys, and consulting engineers, tends to bring about marked variations in the earnings of these groups. But apart from these variations in money wages, these classes of the laboring population constitute no real exception to the law of wages. The normal income of the doctor will be high or low in harmony with the relative utility and scarcity or plenty of the services of this profession or, to use the terms of our law of wages, in harmony with the marginal productivity of the group. So it is with other classes of independent workers.

The more apparent exceptions to the law are exhibited by those groups of workers whose incomes are determined in some material degree by non-economic considerations. Such are those connected with charitable enterprises, members of religious orders, to a less degree teachers and government officials. In many cases the services of these groups are not sold to consumers at prices which measure their marginal utilities to the buyers. There can be no direct relation between the earnings of the workers and the market prices of their products, for the sufficient reason that the prices are not determined by market considerations. What value the consumers would place on the services of men engaged in gathering and distributing the mail we can never know until this business becomes competitive and makes its bid for patronage in rivalry with other outlets for the purchasing power of the consuming public. So it is, though to a less degree, with the school system. We can only guess how many teachers would be employed and what money incomes would be returned to them if these things were left to the test of the marginal utility of instruction to the people as a whole. In the absence of such market tests, the money incomes of workers in social, political, and charitable undertakings must be fixed to a certain extent on artificial grounds.

But even here the economic law of wages is by no means without effect. These occupations must compete for workers with opportunities in the general world of business and professions. The rewards they offer cannot be set too far below those afforded by

competing occupations to men and women of the same ability, else they will be unable to obtain the requisite supplies of labor. The periodic complaint over the dearth of teachers, ministers, soldiers in the standing army, and "good" men in all branches of the civil service shows what happens when the wage policy of these non-commercial enterprises is out of line with the general economic conditions of the times. Making allowance for the non-material rewards of such types of labor as we have been considering — the prestige which attaches to them, security of tenure, relief from strain of competition, congenial surroundings and associates, and the like — we shall find that wages in these occupations show rough similarities to the ruling rates fixed by economic forces in comparable lines in the business world.

CHAPTER XXXIX

DIFFERENCES IN WAGES. WAGE PROBLEMS

The principal labor groups. A comprehensive income census of the United States would show returns from labor ranging from a million dollars a year through all possible gradations down to a few hundreds of dollars. These differences in reward may appear to bear little relation to the exertions of their recipients or to the irksomeness of the different jobs. In the steel mills some jobs are so exhausting that "old age at forty" is the recognized fate of those who fill them. These men have been known to work twelve hours a day, seven days a week, for less than enough to support what we know as a decent standard of living, finally ending their lives in asylums for paupers. In the same industry certain officials put in a normal day of about half this length at apparently light and pleasant tasks, live in great comfort on their incomes, and are able to retire with a competence at a fairly early age. These great contrasts in wage rates arise chiefly from the fact that the laboring population is divided into self-contained classes whose division lines are sufficiently rigid to make them almost entirely non-competing in function.

We cannot of course attempt to classify the thousand and one occupations which absorb the productive effort of our millions of working men and women. Nor is it necessary to a general understanding of this problem to pursue our analysis thus far in detail, since most of these different callings are but subdivisions of larger occupational groups which when assembled comprise the productive life of broad social strata. An individual in one of these particular callings can move more or less at will into any one of a number of others within the limits set by the general aptitudes and acquired abilities of the social class of which he is a member.

Many attempts have been made to classify the working population along these broad lines of social stratification. The following classification which follows the lines worked out by an eminent economist is sufficiently exact for our purposes.[1]

In the lowest group we may place the unskilled workers in all fields of occupation. Every factory has some demand for this type of labor. It is to be found on the harvest fields, in construction work on the railways and roads, in all occupations requiring only a strong back and little training. The supply is limited only by the number of adult males, since almost any able bodied man can do the work required. Wages are paid by the hour or the day and are subject to much fluctuation because of unemployment.

The next group contains the manual workers of little skill whose occupations place upon them responsibilities and require qualities of attentiveness and alertness not demanded of the lowest group. The great body of machine tenders in our factories is an example of this type of labor. The motormen on the street railways, the pick miners, the more common sorts of labor in restaurants and hotels are other examples. Some training and a certain modicum of intelligence and moral stamina are necessary to qualify for these occupations. These requirements limit the supply of labor in this group as compared with that in the lowest class. Wages are somewhat higher and unemployment less common.

The skilled workmen, masters of a craft, form a third group of manual laborers. The subdivisions of the building trades — carpenters, lathers, plasterers, plumbers, painters, steam fitters — furnish the best known examples. The four great crafts on the steam railroads may be grouped in this class. Within the factory system a similar place is held by the tool, die, and pattern makers, the cutters in the garment trades, the compositors and typesetters. These are the aristocrats of the manual workers. Their tasks demand, as a pre-condition, a normal mental equipment and a moral capacity for perseverance; elementary school education and years of apprenticeship are then required to complete the

[1] Adapted from Taussig, *Principles of Economics*, Macmillan, 1916, Vol. II, pp. 134–137.

candidates' qualifications. The earnings of this group are much higher than those of the next lower class, often as much as double.

A fourth group may be distinguished, composed of the class of " white collar " workers. The office personnel of our business concerns, the sale forces of our stores, the lower grades of teachers are examples. These callings attract large numbers by reason of their respectability. They require a general education more prolonged than that of the skilled craftsman but a shorter period of technical training. The social superiority of this class in comparison with the third labor group depends more on its closer association with the well-to-do, whose manners it apes, and on its less irksome and disagreeable type of work than on a higher level of earnings. In fact the wages of this lower middle class are often below those of the skilled craftsmen.

The fifth and highest class contains those higher types of workers who are either self-employed or occupy salaried positions of heavy responsibility and much power. The learned professions, the salaried managers of enterprises, the higher public officials, the independent enterprisers, are examples. These are callings requiring native ability and inherent qualities of character which great numbers of men do not possess. Customary schooling is extended through the colleges and higher technical schools. The earnings are subject to wide variation, ranging from a point not much higher than the wage of the skilled craftsman to the millions of the captains of industry. But despite these differences the entire class enjoys an income in excess of current needs, and all its members expect to accumulate wealth on whose income they and their families may sometime subsist without labor.

" The first three groups, including the manual laborers of all kinds, constitute a class by themselves, not only because the gradations of wages are continuous but because their members have the same point of view and the same prejudices. They expect usually to live on their wages, not looking to the accumulation of property or to an income derived from property. There is a common sense of dependence on manual labor and a sense of separation from the well-to-do and possessing class. The last

two groups have similar feelings of solidarity. Even though there are great variations in possessions and income among them, all have the habits and hopes and prejudices of the well-to-do. They share a feeling that manual labor is beneath them, and their garb indicates their freedom from it — no jumpers or overalls. Their hope is for accumulation and investment, and their ambition is primarily to swing themselves into the position of the leisure class. We may thus divide the workers into two great classes of the soft-handed and the hard-handed." [1]

Wages of each group tend toward uniformity. To the extent that hard lines of social stratification separate these groups, preventing a transfer of membership from one to another, their normal wage rates will diverge. Within each class however there is a tendency for the earnings of the different sub-groups or specific callings to become fairly uniform. This is particularly true of the wage rates within the first two classes. The unskilled laborer may today be a digger of ditches, tomorrow a harvester of crops, later a street sweeper, hod carrier, roustabout, janitor. The rank and file of machine operatives are similarly transferable as between different jobs. Their rates of wages are not, as we have noted elsewhere, entirely uniform. Local discrepancies appear even in the lowest rates in the scale, but any marked discrepancies within the limits of a given labor market are recognized as abnormal and temporary.

The other great social groups differ from the two lowest strata in regard to uniformity of wages chiefly in the length of time required to iron out discrepancies which appear. The third group in our classification falls more definitely into separate skilled trades or callings. The workers in this group at any given time cannot shift from one occupation to another with the same ease as the members of the lower groups, for each specific calling requires a fairly high degree of technical training which, once acquired, tends to commit the individual more irrevocably to his chosen occupation. Bricklayers must be content to remain bricklayers, machinists to remain machinists. Hence there is no ready trans-

[1] Taussig, *op. cit.* Vol. II, p. 137.

fer of the existing workers such as otherwise would occur if the relatively high earnings of one of these callings made it temporarily more attractive than the others. But the distribution of the labor supply which tends to remove discrepancies of this sort is but little delayed. For the skilled crafts call for very little preliminary education of a general nature, and hence the yearly supply of newcomers qualified to learn the trades, of youths who annually assume the duty of self-support, is relatively large. These new workers tend to equalize the numbers in the different trades in proportion to the demand for their services, by applying in increased volume for admission to the trade which for the time being is unusually attractive.

Similar forces may be observed at work within the two upper classes of the working population. The great mass of the workers in our fourth classification — the subordinate "white collar" employees, clerks, stenographers, bookkeepers, and the like — though it falls into a large number of slightly dissimilar occupations, is rewarded with money incomes which vary within fairly narrow limits. The workers in this group can shift from one specific calling to another with greater ease than can the skilled craftsmen, although a longer preliminary period of education is required to equip the youth of the country for this type of service in industry. Commercial schools, "business colleges," night schools, correspondence and extension courses bring the rudiments of the specialized knowledge of business routine within the reach of great sections of the middle and lower middle classes. One of the impressive results of the modern educational system is the tremendous increase in the proportions of young men and women who elect to seek their livelihood in the "white collar" group of workers. These new increments of labor seek out the lines of occupation which offer the highest rewards, thus tending toward an equalization of the membership of the various sub-groups of this class in proportion to their respective demands.

Least uniformity in the highest group. In the highest class uniformity of earnings is least marked. Here not only is the individual worker committed permanently to his chosen profes-

sion by reason of the long and expensive preliminary training required to master a new one, but the annual influx of new workers is small as compared with the lower social groups. Family tradition and other matters of sentiment play a relatively important part in determining the distribution of labor among the professions. Innate differences of native ability and aptitude also exert their influence. Adjustments of numbers to differences in earnings are longer delayed and less automatic. Yet one who observes the succeeding generations of college men at the business of choosing their careers cannot be unaware that the recruiting of new workers in the learned professions and in lines of activity which are comparable in the degree of preliminary training they require is much influenced by the matter of relative earnings.

The independent enterprisers of the business world, who belong to this general social group, are to some extent in a class apart. Here success depends more upon inborn qualifications and less upon training. At their higher points business profits are subject but little to the levelling tendency of competition among workers which conduces to uniformity of earnings between other lines of occupation. This is much less true of the lower levels of business profits. The salaried employee of the highest grade — the young lawyer in the employ of a firm of attorneys, the sales manager, production manager, consulting engineer, industrial chemist — has ever before him the opportunity to venture forth as an independent enterpriser in his chosen calling. Conversely the independent business man may give up the struggle for profits and drop into the more secure position of salaried employee almost at will. These transfers from one position to the other occur when business profits in their lower ranges and the scales of salaries for workers of commensurate ability diverge in any marked degree.

Wages and the contrast of social classes. The great differences in wages which we observe in actual life mark the contrast in economic position of the several social classes. A large railway system usually gives within itself a picture of the whole range of social classes from top to bottom. There are the president with a salary of fifty thousand a year or more, the managers of various

departments with perhaps half this money income, the skilled workers on the trains and in the repair shops with average earnings of about twenty-five hundred dollars a year, the office workers with much the same rate of income, the maintenance of way men receiving little more than a thousand a year, or one-fiftieth of the president's salary.

The cause of discrepancies between wages of the different classes. It is often remarked that the greater differences in wages merely indicate difference in the social importance of the functions performed by the various social classes. As a type of social service, the task of driving a giant engine on a modern railway system appears to be more important than the pick and shovel work of the section gang. The work of the architect who plans and supervises the erection of a beautiful building seems more useful to society than the driving of nails by the carpenter or the lifting of burdens by the hod carrier. That the function of the surgeon is of greater social significance than that of the hospital's janitor appears to be clear beyond all question. This "common sense" view of the matter is often advanced by those who have faith in the inherent justice of present social arrangements as proof that, despite great inequalities, men after all are rewarded in proportion to their deserts.

But we must be on our guard against drawing from such comparisons of individual workers inferences as to the relative importance of functions. Each function is indispensable to production. Imagine a railway system manned entirely by presidents and superintendents, without the labor of engineers, firemen, and trainmen; or a hospital staff composed solely of surgeons with no nurses, orderlies, cooks, or caretakers. When we try to evaluate the different productive functions in their entirety we are on dangerous ground, just as when one attempts to determine the relative importance of land, labor, and capital as separate factors of production.

As a matter of fact, whatever may be the relative importance of the several social classes, it is not this which determines the rate of wages received by the members of each, any more than

their relative social importance determines the prices of bread and diamonds. Here, exactly as in the case of prices in general, we are dealing with marginal utility. In the world as it is, the basic cause of these great contrasts in wage incomes is the distribution of the working population among the various social levels. The economic forces in the labor market do not set values on functions but on individuals; it is the indispensability of the individual which determines his share in the social income. This indispensability is determined in turn by the number of available substitutes for the individual; it is the marginal worth of labor groups that counts. The railway president commands a high price because the number of men capable of performing his function is so very small. The hod carrier's wage is small because he is only one of a vast number of available units.

Permanence of the cleavage between social classes. That great differences in wages persist through succeeding generations is evidence that some obstacle prevents the free flow of labor from one social class to another. If the youth of each generation were entirely free to choose their occupations there would be an immense increase in the number of grand opera singers and movie heroes and a great dearth of farm laborers. Wages at their higher levels would fall and that class which now occupies the lowest station would rise to the top of the social scale.

Wages in different occupational lines would not be entirely equalized, even if there were no barriers between the social classes. The distribution of labor among these different occupations would not reach the same proportions to the demand for their respective services. Some jobs would prove more attractive to the average man than others; perhaps because of the more cleanly work involved, the pleasanter surroundings, the greater independence; perhaps because of differences in the length of time taken to master the details of the job. The exhausting task, the dirty and unpleasant job, the one which involved risk to life or health would be shunned by the workers; and a flood of labor would flow into the more attractive lines of occupation. Coal miners, being relatively scarce, might well get more than managers; and ditch

diggers, for the same reason, more than bank presidents. But these differences in wages would merely serve to compensate for the greater irksomeness of some tasks as compared with others. In other words, given perfect mobility of labor, wages would be equated to the *real costs* of the different jobs, the human suffering, pain, and sacrifice which they entailed. The law controlling these differences in wages would be the same law as now operates; namely, the law of marginal indispensability. But today some barrier to the flow of labor from one social class to another breaks this correspondence between differences in wages and differences in labor cost; indeed, as things go in the modern world, it is most frequently true that the man who works in dismal surroundings at the most exhausting toil is the poorest paid of all.

Nature and nurture theories. Speaking broadly, there are two contrasting explanations offered for the persistence of social stratification through the ages. One group of writers holds the view that conditions of birth and the heritage from his ancestors determine the rank which a man will take among his fellow men. Those in the humblest station are predestined at birth to occupy this position in the lowest income group. Those who rise to higher levels do so by virtue of inborn characteristics which can neither be produced intentionally nor balked of their reward when present. With these writers heredity is the controlling factor in the destiny of men; the present stratification of society is a natural order determined by nature's laws.

In contrast with this view stand the environmentalists. This group asserts that man's social destiny is determined after birth by the forces which play upon his plastic consciousness during his early years. In their opinion, the children of wage earners fail to rise above their father's station in life because of the stunting of mind and body and the impoverishment of will which they suffer from the unwholesome environment of their childhood. Able men are born in about equal proportions in all ranks of society, but those in the lowest ranks are the victims of physical, moral, and intellectual malnutrition. They are denied opportunities for self-development — opportunities which are showered in

abundance upon the children of the rich, regardless of intrinsic merit.

The conflict between the nature and nurture theories lies outside our field; but we should note that the question is not to be set at rest by the superficial evidence usually advanced in favor of one position or the other. The environmentalists do not deny that men differ greatly in inherent ability, nor that these differences are present at birth. The question however is this: does nature distribute these favors of natural capacity with regard to the artificial lines of social stratification — lines determined primarily by the distribution of wealth? Among a given hundred thousand children, one rare genius may appear, a very few will be born with great innate talent, a somewhat larger number with ability beyond the ordinary, the vast majority will be mediocre, and a small residuum will be composed of mental, moral, and physical weaklings. Do these proportions hold good for all income groups in society, or is the proportion of genius and talent higher among children born of well-to-do parents than among those in the lower labor classes? This will always be a moot question until it is possible to separate the two factors, heredity and environment, and observe the operation of each alone. Our institution of inheritance is alone enough to prevent any such experiment. The children of the rich may or may not inherit the abilities of their father, but we know that they do inherit their fathers' wealth, that a favorable environment is constructed for them, that they start the battle of life with a handicap in their favor. Until all children are treated equally without regard to parentage and equally provided with opportunity for self-development, it will be impossible to demonstrate conclusively which of these two factors plays the major rôle in the stratification of society.

Social classes chiefly self-perpetuating. Whatever the ultimate cause of the stratification of society into rigid non-competing social groups, there can be no doubt that the stratification exists and that the different social classes are recruited chiefly from within. Their numbers do not expand and shrink in response to

increases or decreases in the rate of wages which represents the average earnings of their various sub-groups. Striking cases can be cited of the rise of individuals to positions of prominence and wealth from the humblest beginnings; but the class from which these men sprang remains, as a class, what it is and has been — the most numerous and the poorest paid social group. As long as this is true, differences in wages will bear no relation to the real costs of the workers.

Wage problems. Real wages. In any analysis of wages intended to show the comparative well being of the laborers at different times or places, we must be careful to take account of differences in the level of prices. Our discussion of wages thus far has been devoted to money wages. But it is evident that high wages may be neutralized by high prices and *vice versa*. To measure the well being of the laboring class, we must express wages in terms of the quantity of goods which the money will buy. When thus expressed, we use the term *real wages*, as contrasted with money wages. The adjustment of money wages at different times or in different places to a common price level is easily made by means of an index number of prices, either of prices in general or of the prices of the most important commodities and services entering into the typical working man's budget.

Long run changes in the prosperity of labor. For society as a whole, improvement in real income can be due only to some general development affecting the physical productiveness of society's economic system. But it does not necessarily follow that an increase in the total income of society will be evenly spread over the several shares in distribution. It is quite possible that the complex of forces which bring about increased general prosperity may so operate as either to increase or decrease the fraction of the whole which goes to labor.

For example, let us suppose that the number of propertyless people doubles, while the land area, the supply of capital, and the arts of industry remain unchanged. The total product of the community will certainly be larger than it was before, for each additional member of the working population will have brought a

pair of hands with which to aid the productive efforts of the group. But the marginal product of the laborers will be smaller now than it was before their number increased. The productive factors which coöperate with human labor have remained fixed in quantity. There will be less land and less capital for each man to work with. Those laborers who apply their labor directly to the exploitation of natural resources must now work on less fertile lands, less productive mines and forests, or else upon an intensive margin of cultivation which has been forced beyond its former level. In either case, as we know from our study of rent, the principle of diminishing returns will have reduced the product attributable to the individual laborer. In the factories, on the railroads, in other branches of industrial activity, a similar occurrence will have taken place. The additional members of the industrial workers will have been compelled to divide with their fellows the use of the fixed supply of capital instruments; each will receive less effective coöperation from these instruments, and hence their marginal productivity will have fallen.

We cannot say that the aggregate income of the working population will be smaller than it was before. It is more probable that this aggregate income will have increased or, to state it otherwise, that the total share in the social income going to labor is larger in size than it was before the expansion occurred in the population. This depends on how rapidly the marginal productivity of labor falls as the number of workers increases. But after all the size of the aggregate income is not the important thing. What we want to know is how the real income of the average workingman compares now with the average under the earlier conditions. And it is obvious that this average real income has declined as a result of the forces which have caused the total social product to expand. Under the conditions assumed, it is apparent that the owners of land and capital will not have shared the fate which has befallen the laboring population. When the intensive margin of cultivation is advanced and the extensive margin broadened to include lands and other natural agents of inferior productiveness, we know that the rent of land rises. When each unit of capital

instruments is used more intensively, its productive effectiveness is increased. The marginal productivity of capital throughout the whole field of industry will have risen, and the suppliers of loan funds will reap the benefit of this change in the form of a higher rate of interest.

We reach the conclusion that the change which we have assumed would have the effect of increasing the average real income of landlords and capitalists and of decreasing that of the workingmen. By the same line of argument it could be shown that an increase of land and capital, while the laboring population remained stationary, would tend (1) to increase the total social income and (2) to decrease the average real income of landlords and capitalists, while increasing that of the workingmen.

The law of proportionality. The foregoing hypothetical analysis brings to light a principle of universal truth with regard to the relative economic prosperity of the different income receivers in society, which principle may be formulated as follows: *any social change which reduces the ratio which one factor or combination of factors bears to those coöperating with it will tend to raise the average income of that factor or combination of factors as compared with the others.* If population increases faster than land and capital, real wages will fall in proportion to interest and rent. If capital accumulates faster than the increase in population and of natural resources, wages and rent will rise in proportion to interest. If the available supply of land and natural agents expands out of proportion to labor and capital, the rent income will shrink as compared with wages and interest. This generalization we shall call the *law of proportionality.*

Historical changes in the rate of wages. Differences in the general welfare of the industrial populations of different countries, or of the same country at different times, are to be explained in the light of the basic determinant of wages; that is, the marginal productivity of labor. There have been epoch-making changes of this character in the recent history of the world. The period since the beginning of the Industrial Revolution has been one of revolutionizing discovery of new sources of power, rapid accumu-

lation of capital, progress of invention, improvement in transportation, perfection of the science of business organization and management. In 1801, the total population of England and Ireland was approximately 16,000,000; by 1901 it had risen to 41,500,000. At the earlier date the economic condition of the mass of people in England was one of appalling poverty and misery; during the century, despite the rapid multiplication of people, real wages rose steadily. The following table traces this rise of average real wages from 1830 to 1900, using the average of 1900 as a base.[1]

Year	*Index of Real Wages*
1830	45
1840	50
1850	50
1860	55
1870	60
1880	70
1890	84
1900	100

The average real income of labor in the United States has always been higher and is now higher than in older countries less abundantly supplied with natural resources. The steadily rising wage income of the first three quarters of the nineteenth century must be attributed principally to our expanding frontier, which opened up seemingly inexhaustible supplies of land. This increase in coöperating factors permitted an unparalleled expansion of numbers — from 3,929,000 in 1790 to 38,558,000 in 1870 — while allowing the marginal productivity of labor steadily to rise.

A change in basic factors occurred during the last three decades of the century. During this period the frontier was gradually extinguished, but there was a rapid increase of capital, a continuous progress of invention, an extension and perfection of the transportation system, and a revolution in industrial organization and the science of management. The increasing factors were primarily capital and entrepreneurial ability. Expansion of these factors in other countries, together with some discovery and

[1] From A. L. Bowley, *Statistical Studies.*

exploitation of formerly isolated natural resources, contributed through the extension of our foreign commerce to improve the economic status of our laboring population. During these thirty years, the number of people increased from 38,558,000 in 1870 to 75,995,000 in 1900. The rise of average real wages is shown in the following table in which the average wage of 1890 is used as a base.[1]

Year	*Index of Real Wages*
1870	68.7
1875	72.5
1880	82.8
1885	98.2
1890	100.0
1895	102.0
1900	104.5

Since 1900 another change may be noticed in the progress of the average wage income. For the world as a whole, population continued to increase, but free supplies of land and natural resources disappeared, and, though capital increased, its expansion was not rapid enough to sustain the former rise of wage rates for the increasing population. There is little doubt that the average wage income for industrial nations rose very slightly, if at all, and that in some countries the real income of the average working man began to fall. In the United States population increased from 75,995,000 in 1900 to 105,711,000 in 1920. The national product per capita for the entire population expanded by about 28 per cent; but wage labor did not share in this increasing prosperity. This fact is disclosed by the following table, in which the average wage from 1890 to 1900 is used as a base.[2]

Year	*Index to Real Wages*
1900	101.6
1905	98.3
1910	87.8
1915	86.6
1920	89.0
1923	95.0

[1] From Adams and Summer, *Labor Problems*, p. 514.

[2] Paul H. Douglas, "Movement of Wages and the Future of Prices," *Annals of the Academy of Political Science*, 1925.

This, it should be noted, is a statement of the trend of the average wage income. Many specific labor groups rose in the scale of prosperity during this period; others fell. But for wage labor as a whole, the trend was downward, and this at a time when the total national product was increasing faster than the population. The period is too short and too much upset by violent swings of the business cycle and by war conditions to permit us to draw definite inferences from these data. But the evidence points to the conclusion that the balance has turned against labor and in favor of the other productive factors. Certain it is that, with the preemption of all available land areas, the number of workers in the world cannot continue to increase at the present rate of speed without a decline in the average wage income, unless the balance is continuously rectified by increases in the amount and the efficiency of capital. Labor as a whole therefore has the most vital interest in all those forces which support and extend the capital equipment of modern society.

Law of proportionality applied to specific labor groups. We have been speaking of the real wages of labor as a whole. But the law of proportionality applies in detail to the relative real wages of the different classes into which the entire laboring population is divided. If each particular labor group consumed its own product, its material welfare would be increased by any force or combination of forces which multiplied the physical productivity of its members. But this we know is not the case in modern society; the standard of living of one of our occupational groups is made up to but slight extent of the products of its own toil. The workers in the shoe factories are much more concerned about the supplies of foodstuffs, clothing, dwellings, coal, and other things, than they are about the supply of shoes. Each group, in short, is much more concerned with the productivity of other groups than it is with its own.

This feature of our exchange economy has far-reaching effects which are not always favorable to general social welfare. It means that it is often to the immediate economic advantage of each separate labor group to restrict its membership and raise the

value of its marginal product. Many devices are adopted to produce this result. When the workers in a given occupation are organized into trade unions and are able to prevent those who are not union members from engaging in the occupation, they are tempted to adopt restrictive practices which artificially limit their numbers and, in an equally artificial manner, increase the numbers in other lines of activity. This will raise their marginal productivity, make their product relatively scarce as compared with other products, and increase their real wages. Much the same result is obtained if the group in question adopts "make work" policies which set a limit to the output of the average man in a given time. This may be done by deliberately lowering the rate of output below the normal capacity of the workers, opposing labor-saving machinery, destroying the effectiveness of waste-reducing devices, and in many other ways. These are policies not peculiar to the workers; they are the stock in trade of all monopolists. But they are practices which in recent years have been most commonly associated with labor, having been forced upon the attention of the public through the rapid growth of trade unions in numbers and in power.

It is customary to condemn such practices as these on the ground that they are destructive of the welfare of the very men who adopt them. But this is the opposite of the truth. It is not difficult to prove that society as a whole is always injured by any force which artificially reduces the productivity of any of its classes of workers. But no single group of workers is interested in the economic welfare of society as a whole; it is anxious about its own rate of real wages. Now the fact is that each labor group is injured by such restrictive policies as we are discussing only when they are practiced by others; each is benefited by the same policies when practiced by its own members. To be sure there are limits beyond which this practice cannot be pressed by the group without causing it to react indirectly against its members. If the services of the group become too scarce and of too high value, substitutes will be found which may displace them entirely. And of course if all groups of workers begin to employ such devices

at the same time, they will checkmate each other, and the outcome will be general impoverishment for everybody concerned. Nevertheless, restrictive devices not only result naturally from the existing arrangements of the social order, but are often sensible and "practical" from the standpoint of the group in question, however injurious they may be to the general social welfare.

The standard of living theory of wages. If it can be said that the workers themselves hold any particular theory of wages, it is that wages either are or ought to be determined by the laborer's standard of living. It is not to be expected of course that any theory or explanation which arises from the thinking of the mass of men in industry should have been worked out with scientific accuracy. For the most part, the wage theory of the workers lays stress upon what is fair and just rather than upon what is economically possible; in other words, it falls in the domain of ethics rather than in that of economics. With this phase of the question it is not our task to concern ourselves. But the standard of living theory is often advanced as an attempt to explain things as they are, and as such it vies with the economist's theory of marginal productivity and must be considered in any study of wages.

In its most extreme forms the standard of living theory encourages extravagant expenditure on the part of the working classes. For over a century the wage earners have opposed the efforts of social workers to reduce their cost of living by means of frugal purchasing of supplies, cheaper diet, thrifty budgeting of costs and income, and similar devices. It has been their belief that any gains resulting from these expedients would be appropriated by the employers in the form of a lowered wage rate. In the reverse direction many labor leaders and individual members of the rank and file have favored the inclusion in the workers' normal standard of living of expensive items of consumption. One of the more dramatic of the American labor unions had a slogan which ran: "A dress suit and a bottle of champagne for each worker's dinner table." The idea here is that such extravagances, as soon as they become habitual, will play their

part in determining the rate of wages. Underlying this whole theory, we usually find unexpressed the belief that the employer's remuneration of his hired employee is determined either autocratically or at any rate arbitrarily and that it tends to gravitate toward some cost-of-subsistence minimum where its decline is arrested either by the force of sentiment — public sentiment or the employer's compassion — or by the sheer impossibility of obtaining a supply of workers for a smaller wage.

The true effect of the standard of living. The fallacy in the crudest of these corollaries of the standard of living theory is so obvious that we need not stop to discuss it. But there is an element of truth in the law itself which should be understood by students of economics, especially in these times when it is popular to attempt to fix wage rates with reference to the cost of living. There are two ways in which a high standard of living may be effective in raising the wage rate or preventing its decline. In the first place a firmly entrenched standard of living may operate to increase the bargaining power of the workers and enable them to resist a reduction of wages or to win an advance. To have this effect the worker must be so firmly attached to his standard that he will refuse to work for a wage which involves a degradation of this traditional mode of life. When this is true, the standard of living sets the lower limit to the range of wage bargains; a gradually rising standard will elevate this lower limit and cause the actual rate to approach more closely and adhere more permanently to the point determined by marginal productivity. In view of what has already been learned on this subject, it need scarcely be pointed out that the wage cannot be maintained by these means at a point above the marginal worth of the workers for any considerable length of time. Nevertheless, there are on record enough cases of practical success in winning increases in wages to lend empirical support to the argument that the actual wage rate may be influenced by the standard of living.

In the second place the standard of living must be viewed as one of the forces determining the marginal productivity of the workers, in which character it plays a part in the complex of eco-

nomic forces which fix the normal rate of wages. The reason for this is two-fold: the standard of living affects the productivity of the individual workmen, and in a fundamental way it also exerts control over their numbers. It is common knowledge that a close correlation exists between the productive efficiency of the worker and those phases of his standard which consist of physical maintenance. More nourishing food, better clothing, better housing conditions improve the physical and mental alertness of the workingman and resist the destructive effects of fatigue. We are also beginning to see that less tangible elements in his mode of life have a similar effect upon the worker. Leisure for intellectual and spiritual development, freedom from worry and anxiety, a hopeful prospect of social advancement for his children, all contribute to the worker's effectiveness as a producer. The argument in support of the "economy of high wages" rests upon such considerations as these, claiming that, when once the wage rate of the low-paid worker has been raised, he automatically becomes worth the higher wage. We need not push our analysis of this phase of the standard of living theory very far in order to appreciate its truth and also to grasp its limitations.

But the wage rate of an individual member of a labor group is governed, not solely by his own physical fitness and his own industry, but also by the bare fact of numbers within his group. The standard of living operates over long-run periods to set a limit upon the numbers of workers in the different occupational groups, as well as to the total number who seek employment in industry. Viewed from this broad social standpoint, the standard of living is defined as "that quantum of physical and psychical enjoyments which the individual demands in preference to marriage and the procreation of children after marriage." In this day of deliberate postponement of marriage and equally intentional control of the birth rate, the importance of this factor as a check upon the growth of population will be granted without much argument. By advancing its standard of life, a given labor group not subject to infiltration from other classes of workers may make itself progressively a *scarce* factor in industry, raising its marginal produc-

tivity over a long-run period of time and advancing its rate of real wages.

For certain wage-earning groups this implication of the standard of living theory has come to have very practical meaning. The higher grades of workers began long ago to give evidence of the influence of their standard upon their rate of increase. As a result it has become a commonplace among social students to attribute to this influence some of the credit for that rise of families in the social scale which has characterized society since the Industrial Revolution. Unfortunately the more necessitous groups, whose poverty is due to their excessive numbers and a consequently low marginal productivity, are least affected by this factor. It is among the classes which are already prosperous that we find the greatest inclination to elevate the standard of living in preference to an increase in the birth rate. As a result, the principal influence of the standard of living, viewed as a check upon numbers, has been in the direction of maintaining existing high wages and, over long-run periods, slowly advancing them, rather than of raising the low wage rates.

Elevation of the wage rate through action of the employer. The standard of living has come to hold an important place both in public policy and in the wage policies of certain individual employers. For example the wage policy of the Standard Oil Company of New Jersey contains as one element the rule that the wage rate must not fall below a reasonable standard of living. The Filene Company has a similar policy, as have also the Endicott-Johnson Company, the Dennison Company, and many other large employers of labor.

It must be admitted that individual employers are sometimes at liberty to set their own wage rates above the point elsewhere determined by economic law. Prosperous concerns operating on a large scale, employers who possess unusual ability as managers, and above all those enterprisers who have a monopoly or something approaching monopoly in their respective lines of industry may maintain the marginal productivity of the class of workers they employ above the general level of this class in the business

world. This makes it possible for them to pay wage rates above the normal or average, and certain considerations sometimes make it desirable for them to do so. We may leave out of account philanthropic and compassionate motives, which undoubtedly have some influence, but which operate only in exceptional cases in the world of business. Of much more importance are the considerations which make it profitable for employers with especially prosperous or quasi-monopolistic enterprises to adopt this line of wage policy.

Some years ago the Ford Motor Company decided to enlarge the scale of its operation to include the production of mechanical parts which it formerly had bought ready-made from other manufacturers. This decision necessitated the employment of large numbers of skilled mechanics — not an easy task at a time when little unemployment existed in the labor market. Having constructed its plant, the company announced through the newspapers a rather startling minimum wage policy. The most humble employee in the plant was to receive no less than five dollars a day, while the more skilled workers were offered proportionately higher wages. Since these rates were very much in excess of the normal earnings of the workers at the time, this generous offer quite naturally received nation-wide publicity. As a result the best workers in other industries flocked to Detroit to seek employment in the new concern, and the Ford Company succeeded in a very short time in getting together a labor organization sufficiently above the average in productive capacity to be worth the higher wages.

Most employers who fix their wage rates above the normal earnings of the group in question, equating them to some hypothetical standard of living, will frankly acknowledge that they do so because it is "good business." The incident which has been related indicates one way in which the employer may gain from such a policy. He obtains a picked lot of workers, the "cream" of the supply. The policy also makes for permanence of employment, reducing what is called the "labor turnover"; *i.e.*, the habit of workers of drifting in and out of jobs, often in response

to whim or caprice. Since there is some expense connected with the hiring of new workers and still more involved in breaking them into the routine of the plant, the stability of a labor force contributes to the profits of the enterprise. A contented and stable group of seasoned employees also makes for discipline in the shop and for the smooth functioning of the plant. These are but suggestions of the ways in which an attractive wage policy may be "good business." It is to be noted that these results do not follow from the fact that wages are high; but from the fact that they are higher in one enterprise than in industry generally. The industries offering the high wages do not pay their workmen more than the value of their marginal product, but they employ devices which raise this marginal value above the ordinary level of the business world. This then is no real qualification of the law of wages.

Minimum wage laws. The attempt to fix wages by statute over large geographical areas and many competing industries is another matter. For example some dozen states in the United States have minimum wage laws applying to women. The economic law of wages does not deny that wage rates can be raised to an artificial level by statute, providing the statute can be enforced. It is conceivable that such statutes may materially raise the rate of wages for a given group of workers — women workers let us say, or coal miners — over the entire field of industry. That employers can be made to pay minimum wages has been demonstrated by the activities of trade unions. With the whole power of government behind such a policy it may be made impossible for anyone to hire labor at less than an unusually high money wage.

But though minimum wages may be fixed at a high level by law, there can be no guarantee that all the workers will find employment at the legal rates. If employers are prevented from equating money wages to the marginal productivity of the labor group through the device of lowering the wage rate, there remains another method of bringing about this equivalence; namely, through a reduction in the number of workmen employed. As we

have seen, the value of the marginal man's product rises with each decrease in the labor group (conditions of demand remaining constant) and falls with each increase. After wages have been settled by competition at some point not far removed from the economic worth of the marginal workers, any artificial elevation of the rate will force employers to discharge some of their workmen, this being the only available method of avoiding loss. If pressed too far therefore the minimum wage law will create the problem of unemployment. This must be so in an industrial system founded on the principle of free enterprise, where employment is voluntary on the part of the employer.

Hence the device of the minimum wage is clearly ineffective to produce a far-reaching and appreciable improvement in the condition of the workers. If employed for the benefit of a restricted labor group, its net result after the wage rate has been forced above the competitive level will be to precipitate a portion of these workers into a lower group where their condition will be worse than before. On the other hand, if the policy is made all-embracing and attempts to advance all wage rates simultaneously, the real wages of the different labor classes will not be affected by it. Regardless of the manipulation of their money wages, if they produce the same quantities of their respective goods and services as before the passage of the law, the flow of real income into the hands of the working population in general will not have been altered.

These conclusions do not go to prove that legal minimum wages have never turned out to the permanent benefit of laborers. Indeed a study of the history of this line of public policy proves the contrary to have been the case. It is clearly shown on the record that the minimum wage laws of England have brought about a considerable improvement in the real income of certain labor groups without causing unemployment, and this is also found to be the case with the laws of our states relating to the labor of women. These facts of experience argue either or both of two things. First it may be that the competitive wage in operation before the passage of the law was lower than the economic wage

or, in other words, that the bargaining weakness of the workers had been exploited to the end that they were paid less than they were worth. In such a case, as we have already seen, the actual wage rate can be raised by any device which strengthens bargaining power, and this device might well be put into the worker's hands through government enactment. Secondly, workers of low productivity may have shown an increase in efficiency in response to better living conditions and the greater contentment which followed upon a rise in their standard of life. This factor too has been discussed in the preceding pages. The initial rise of wages which produces this improvement in productive efficiency may come through legal enactment, but if the ultimate result is a rise in the marginal productivity of the group, it is evident that the higher wage rate may become permanent, and that with no reduction in the numbers employed.

Whether the minimum wage will prove a real boon to labor or an ultimate evil depends then upon the caution with which the policy is applied and the extent of improvement which is attempted through it. Examination of the laws of the United States and England which have succeeded in raising wages without creating unemployment show that these limitations have been carefully observed. The standard method of determining the legal minimum has been to take the wage policy of the more generous employers in a given occupational group as a standard. The statute in most cases merely makes this policy general throughout the industry, enforcing it against those employers who have been prone to use their bargaining power to exploit the workers. Where the higher wage policy has been supported by the fact that they have had a better type of labor than the average of the group, the minimum wage law has also tended to bring the average up to this level of productivity through its effect upon the efficiency of the workers. Such experiences are quite in harmony with the economic law of wages and leave unchanged the implication of that law — that any attempt through artificial means greatly to increase the wage rate will rebound upon the workers in the form of the more serious evil of unemployment.

CHAPTER XL

PROFITS

The word which stands as the title of this chapter is one of the most commonplace expressions in everyday speech. Like most everyday expressions, it is used in a variety of meanings, none of which is entirely exact or definitive. To give precision to the concept of profits is no easy matter; yet, difficult though it be, this is what we must now proceed to do, lest we become involved in hopeless confusion at the very outset of our study.

An example of profits. As an aid to the first step in our analysis, let us consider the following hypothetical income account of an independent business man.

INCOME ACCOUNT OF JOHN CARPENTER, COAL DEALER
Jan. 1, 1925 to Dec. 31, 1925

Income		*Outgo*	
10,000 tons soft coal @ $9.00	$ 90,000	Cost of soft coal @ $6.75	$ 67,500
2,500 tons anthracite @ $15	37,500	Cost of anthracite @ $12.75	31,875
600 cords wood @ $10	6,000	Cost of wood @ $8.00	4,800
		Interest on capital $40,000 @ 6%	2,400
		Rent	1,000
		Wages	9,250
		Heat, light, taxes	1,100
		Attorney's fees	100
Total income	$133,500	Office supplies	450
Total outgo	119,075	Depreciation	600
Net income	$ 14,425	Total outgo	$119,075

The reader will have no difficulty in recognizing all the items in the left hand column as income to the owner from his business, and it is equally clear that the items of the other column are costs of the business. The principal elements of business costs are those dis-

tributive shares which we have already studied — economic rent, interest on capital, and wages — and in addition prices of materials and miscellaneous payments for necessary services, most of which if analyzed could in their turn be resolved into economic rent, interest, and wages.

The difference between the total income of the business and its total costs, $14,425 in the above account, will be regarded by the coal dealer, in harmony with what is perhaps the most customary connotation of the term, as his "profits," the fundamental idea being that profits are what is left to the owner out of his business income after the deduction of business costs. But before we can feel confidence in this concept we must subject it to some further analysis.

Rent and interest are costs. In the hypothetical account before us, it may safely be assumed that the capital used in the business is borrowed and that the $1,000 of rent is paid for the use of land belonging to another person. But now let us change the supposition by assuming that, all other circumstances being as before, the land and capital used in the business belong to the owner himself. A crude accounting system, while showing the same total income as above, might show total costs of only $115,675 and a net income of $17,825. This would clearly be the result of omitting the items of interest and rent from the costs, justified by the admitted fact that no such payments were made to anyone. The resulting "net" income however is clearly not all to be ascribed to the operation of the business, since the owner might close up the business entirely and still derive $1,000 from the rent of the land and $2,400 in interest on his capital. Accurate accounting therefore requires that rent and interest be charged as business costs even when they accrue to the owner of the business himself.

If the reader finds difficulty in reconciling our present inclusion of rent among business costs with our previous conclusion (at the close of Chapter XXXIV) that economic rent is not a cost of production, he will do well to refer once more to the discussion at that point and then to recall that there are two viewpoints, the social and the individual. From the point of view of society as a whole,

economic rent is no cost, since it is a purely differential gain which does not affect the price of the product. To the individual business man, whether owner or tenant of the land used in his business, the economic rent must be regarded as a cost of his business. Since he could continue to enjoy it even if he retired from business, it clearly must be deducted in determining the gain from the business as such. From the point of view of his own business, rent is therefore a cost. It still remains true, as previously pointed out, that economic rent does not affect the price of the business man's product.

Wages of management a cost. But now arises the question, should not remuneration for the business man's own efforts be treated in the same way as interest and rent? Let us suppose that our hypothetical coal dealer decides to retire from the active management of his business and that he hires a manager to perform this function for an annual salary of $8,000. Now the manager's salary will certainly have to be included in the costs of the business, increasing the total costs to $127,075 and reducing the net income to $6,425. Yet the business is just as successful as before, the difference depending merely on whether the managerial function is performed by the owner himself or by an employee. This clearly impels us to the conclusion that a reasonable remuneration for the services of management must be charged as a cost in determining the true profits of any business, whether the manager is a hired worker or the owner himself. The real profits of the business in our illustration are not $14,425, as would commonly be assumed, but $6,425. It is this amount which the owner derives from his position as an independent enterpriser.

If the reader perchance finds difficulty in accepting this conclusion, possibly at variance with preconceived notions, he may be helped by reverting to the analogy of rent and interest. The coal dealer in our illustration, assuming that he owned the land and capital employed, could continue to receive the rent and interest thereon even though he gave up the business. It is clear enough that the $17,825 of crude net income included $3,400 of rent and interest, which were no part of the real profits of the business.

But in exactly the same way this man could, by giving up his business, find employment and obtain a salary for performing services similar to those which he now devotes to his own business. If his worth is such that he could command a salary of $8,000 from an employer, it is clear that he is sacrificing this salary when he devotes his energies to his own business and that $8,000 of the apparent net income of the business is not a true business profit, but is really of the nature of wages.

Rent, interest, and wages all stand on the same footing. It is perfectly obvious, when all these are actually paid out for the services of rented land, borrowed capital, and a hired manager, that they are business costs; here rent, interest, and wages are explicit. But they are no less present and no less costs of business when they are implicit; that is, when they are to be ascribed to the owner of the business himself for his own land, his own capital and his own labor of management. We shall find the term, *wages of management*, a useful one to designate the remuneration for business management, which is implicit in the case of the man who runs his own business.

It is probably true that this concept of implicit wages of management as a business cost is at variance with the notions of the ordinary man on the street. But the logic of the situation is today rapidly forcing itself upon the attention of business men. The recent extraordinary development of the science and art of cost accounting has led to clearer notions on the part of business men. In particular the income tax has brought the issue down from the realm of academic theory to the bed rock of practical business, since there is no possibility of equitable income tax accounting without making adequate allowance for the services of the owner of a business. As a result it is probable that the great majority of independent business men today recognize an allowance for their own services among their business costs.

What governs the rate of wages of management? As modern industry is organized, the services of the business manager, whether conducting his own business or working for another, are indispensable. There is a market for this type of labor, with a scale

of wages determined in accordance with principles with which the reader is now acquainted. The remuneration of any particular business manager will depend upon his own efficiency as well as upon the general conditions of demand and supply governing the type of labor which he represents. The man who is competent to perform the managerial function is as a general rule in a position to choose between engaging in business for himself or hiring out to an employer. He is presented with a choice between two types of income, the relatively stable and secure salary income offered by the wage contract, and the insecure and fluctuating income from business. We cannot say with assurance how any individual will react to this choice. The cautious and timid very probably would require a reward higher than their normal wage to tempt them from the security of a salaried position; the courageous and those who prize independence and freedom might accept a return smaller than their alternative salary. But on the average and in the long run we cannot believe that enterprisers will be content with a reward which falls below their alternative wage income, any more than that they will forego interest on their capital and rent from their land. And with millions of individuals transferring at will from one status to the other, there is obviously a tendency for wages of management and the contractual wages of that grade of labor to approximate each other in the long run.

The definition of profits. We have now arrived at a clear and exact concept of profits and are in a position to formulate a definition. *Profits are the net income, or the difference between the income and the costs, of a business, the costs including economic rent, interest, and wages of management, either explicit or implicit.* Profits belong of course to the owners of business enterprises. They are the reward of that class — business men, entrepreneurs, enterprisers — for the service of exercising judgment, forecasting the future, and bearing the risks required by modern industry. We must also note that profits are essentially variable and uncertain. In any given enterprise they may be large one year and small the next. Costs may exactly equal income, in which case profits are zero; or costs may even exceed income, resulting in losses instead of

profits. We shall find it an aid to our analysis to regard losses as negative profits, which makes it possible to say that profits are always obtained by subtracting costs from income, even though the former be the greater and the result therefore a negative quantity.

The existence of profits. Of the actual existence of profits, as we have defined the concept, there is evidence a plenty. It is true that there may be difficulty in measuring profits or even in demonstrating their existence in those cases where the business man's income includes implicit rent, interest, or wages of management. Placing a money value upon these items of business cost — particularly the last — is an act of judgment, which of necessity gives some margin of uncertainty as to the amount of the residual quantity, profits. But this difficulty is not always present. Extension of the principle of division of labor has had the effect of disclosing the pure profit income, in specific cases, in contrast with all other types of income. Many a corporation not only makes a separate accounting of rent, interest, wages, and other business costs, but hires at contractual wages all the managerial labor it requires. From the president down, all the managers are salaried employees receiving wages, which differ only in amount, not in essential character, from the wages of the humblest laborer. This practice places wages of management in the category of explicit business costs and makes it possible to earmark profits, if any exist. The United States Steel Corporation for example has at its head a hired laborer who supervises the function of management with the aid of various subordinates, who are likewise hired laborers. Each of our great railroad systems has its salaried president, vice-presidents, and department heads. So it is with the larger banks, trading companies, and manufacturing concerns.

In such cases, any residuum of net income appears clearly in the light of pure profit. If we focus attention on specific industries or specific individuals, we find abundant proof of the existence of this income. During the twenty-year period between 1901 and 1922 the common stockholders of the United States Steel Corporation accumulated profits of over three quarters of a billion dollars quite

in addition to dividends which they received. In one ten-year period the Standard Oil Company of New Jersey paid annual dividends of 20 per cent to its common stockholders and accumulated several hundred millions in profits. Henry Ford began his manufacture of motor cars with an investment of $49,000 and within twenty years had reinvested profits amounting to nearly half a billion dollars. The great fortunes of modern times are based on profits. The railroad industry has produced enough profits to provide gigantic incomes for the Goulds, the Vanderbilts, the Harrimans, and the Hills to the end of time. The banking firms are examples *par excellence* of the recipients of pure profit. The funds they use are but to small extent their own, and on these funds they receive interest; they have little to do with the actual running of enterprises, yet they make tremendous profits.

Losses. Such examples as these however should not be allowed to stand alone. The other side of the picture discloses an impressive record of losses or negative profits.

Statistics show that at the end of each year more than one hundred thousand corporations have nothing left for dividends after payment of business costs. In 1917, a year of exceptional business prosperity, 119,000 showed an aggregate deficit of 629 million dollars. The entire business world is subject to startling reversals of fortune. In 1919 only about two thirds, and in 1921 less than half, of American corporations reported net earnings to the United States Treasury. In 1921, out of a total of 356,897 corporations reporting to the Treasury, 185,158 showed a combined loss of four billion dollars, an amount only a little less than the combined net earnings of successful corporations. This was a bad year, to be sure, but it serves to illustrate the fluctuation of earnings resulting from business risk.

As illustrating similar fluctuations in the fortunes of industrial groups, the following statistics are illuminating. In 1920, 73 per cent of the corporations in the lumber trade reported net earnings; in 1921 this ratio had fallen to 45 per cent. In the metal trades the percentage of profit-making concerns changed from 67 per cent in 1920 to 35 per cent in 1921. Similar changes can be

seen in the income records of individual concerns. In 1917 Swift and Company earned 46.2 per cent on its stock; in 1920, 3.4 per cent; in 1921 this concern showed a deficit of 5.2 per cent on its capitalization. In 1915 the United States Smelting Company earned $11.39 on each share of common stock; in 1919, $14.06; in 1922, nothing; in 1924, 28 cents. The earnings of the Atlantic Refining Company for the years 1916 to 1924 have ranged from $220 per share down to a deficit of some $60 per share.

Further evidence of negative profits is given by the record of bankruptcies. In the best years about one per cent of all the businesses in the country are wiped out by insolvency, and this ratio is more than doubled in years of depression. The following table gives some significant figures.

RECORD OF BANKRUPTCIES IN THE UNITED STATES
1917–1922

Year	*Number of failures*	*Total net loss*
1917	13,029	$ 81,770,000
1918	9,331	69,565,000
1919	5,515	60,188,000
1920	8,463	179,224,000
1921	20,014	309,179,000
1922	21,400	282,203,000

The balance of profit and loss. Our statistical data are not sufficient to permit of anything approaching an exact balancing of total profits against total losses. But we may at least hazard the guess that on the whole business losses are not far short of business profits and that there is therefore little or nothing in the way of a net income going to the whole class of business men as a reward for their particular function of business forecasting or risk bearing. This opinion is doubtless at variance with popular notions. When computing the result of chance in any hazardous undertaking, it is a common failing of human nature "to count the hits and disregard the misses." The casual observer of the results of economic activity in society is so blinded by the glitter of the large prizes that he overlooks the frequency of losses. But for every case of large profit in a given enterprise or for a single individual business man, a case of loss can be cited for another enterprise or individual

which tends to offset the profit in our accounting of the distribution of society's income. The reasonable conclusion is that in the distribution of the social income the class of entrepreneurs as a whole receives at best a surprisingly small share.

Profits a reward for risk-bearing. When profits arise in individual enterprises, they accrue to the owners, not to the managers nor to anyone else who takes an active part in the current operations of the business. When corporations hire their managerial labor at contractual wages, the pure profit, if any exists at the end of the accounting period, is merged in a fund called "undivided profits" destined either to be distributed among the stockholders or to be reinvested in the enterprise with the result that the value of the stock is increased. The position of the common stockholder, who receives the major portion of the profit, differs from that of the preferred stockholder, whose rate of return approximates the market rate of interest, and the position of the latter differs from that of the bondholder, whose share in the profits is nil, chiefly in respect of the amount of risk borne by the different groups. It is impossible to eliminate risk even in the routine operations of the most firmly established enterprises.

In the routine of our economic experience most of us live from day to day. We have always found the essentials of existence ready to hand in the market and confidently expect to find them so tomorrow; our problem is to obtain a steady stream of purchasing power with which to buy the things we need. To cite a familiar example, everyone knows that the supply of oil is definitely limited and that one day there will be no more gasoline obtainable in commercial quantities. But of the millions of motor car owners, how many are engaged in discovering a substitute? World events reported in our newspapers are read by countless numbers of people, but only a handful take thought to winnow out those occurrences which threaten the future economic welfare of ourselves and our children and to lay plans to counteract their effects. When present occurrences, events, tendencies of such far-reaching importance are known, the sensible thing for society to do is to discount them and make timely adjustment. The vast majority of us leave all this

out of account, to be attended to by a self-constituted group of men who possess the requisite foresight, courage, and energy. We may call these men the risk-bearers, if we are careful to keep in mind the broad implications of the term; but their major social function is that of guiding and directing the course of economic development so that future conditions will find us prepared in advance to grapple with them. Those who draw profits from an anticipation of future costs and income are doing something of which the term risk-bearing as ordinarily used is scarcely descriptive. It would be more illuminating of their real service to society to say that they *capitalize their prevision of the future.*

Risk, insurance, and profits. For various reasons the prices which determine income may be out of adjustment with the prices which control costs. We may classify the causes of these maladjustments roughly in two general groups: those which occur as commonplace incidents in routine business operations, and those which have their roots in the dynamic character of our industrial structure. The former affect the profitableness of individual enterprises within an industrial field; the latter exert their effects upon industries in their entirety. The most striking cases of pure profit arise from the broad secular changes which are embraced in the second category of maladjustments. The ordinary risks of business, when they become foreknown, lose their uncertainty and no longer call for the type of service which we associate with the activities of the enterpriser. The development of the practice of insurance makes this clear. One by one many of the risks of modern industrial life have been brought within the compass of mathematical formulae and have had the law of averages applied to them. When it is discovered that a given risk is insurable, the losses which it entails are apportioned among all affected by it. It becomes one of the normal and calculable costs of doing business and plays its part in determining market prices. Thus the business of insurance may be viewed as a continuous encroachment on the domain of the risk bearer and upon the profit income. If all hazards could be handled by the insurance method, most of the maladjustments in the price system

would disappear with corresponding disappearance of the pure profit income. But insurance has not yet been made applicable to changes in processes, in the range of human wants, or in the availability of natural productive agents which arise from the dynamic nature of our social life.

The promoter and his profits. We are now prepared to understand why men receive profits who have nothing to do with the current operations of industry. Consider for example the function of underwriting new ventures discharged by the investment banks. When they believe that timely provision of capital equipment along novel lines promises lucrative returns in the future, these banks guarantee to provide the necessary supplies of present funds, partly out of their own resources, partly by inducing others to take a share in the venture. Through their efforts capital instruments are constructed and assembled, managerial and manual labor is organized, the different parts of a productive establishment are brought into coöperation with each other. If the forecast is correct, this productive unit will be equipped and ready for service at the time it is needed. The bankers will then withdraw and surrender the actual ownership and management of the new enterprise to others. But they will have capitalized the opportunity for profit by retaining for themselves shares of stock in the enterprise, whose dividends absorb a large part of the pure profit which accrues. It often happens that these first comers in the field are compelled to divide their title to the pure profit with other investors upon whom they rely for a part of the required funds. This they may do by offering a rate of return on preferred stock in excess of the normal rate of interest, or by selling common stock at a nominal value, or even by giving it outright as a bonus to the men who venture their funds in the purchase of bonds or other securities. But a large slice they retain for themselves in the form of common stock for which they have made no payment other than the labor of forecasting, planning, balancing possibilities of future income against future costs. Such stock is merely legal title to pure profit if it accrues. The investment bankers may retain it permanently, thus receiving a continuous stream of profit from a successful venture,

or they may sell it to others after the profitableness of the venture has been established and thus acquire outright the capital value of their share of future profits.

We have just described a process known in the business world as *promoting* new industries. The class of professional promoters has been called into existence by the dynamic character of modern industry. They are a group of the shrewdest and most far-sighted of business men, who devote their whole energy to forecasting the currents of change at work and placing themselves in position to profit by an excess of future income from the sale of goods and services over the expenditures entailed in costs of production. A new railroad is built across a stretch of sparsely settled prairie at a cost of scores of millions. Its promoters are predicting a thickening of population in the region, a shifting of the geographical location of industry, a new relation between producing and consuming markets. If these changes occur, the promoters will have rendered society a service by providing capital instruments of immense worth in the new setting, and they will profit accordingly. Another foresees that the need of a great nation for meat products cannot long be satisfied piecemeal from local sources of supply. He constructs a plant for the curing and preserving of meats, perfects the technical processes, elaborates the buying and selling mechanism, and profits with the growth of the conditions which he has foreseen. Illustrations of the promoter's activity could be multiplied indefinitely, for they lie ready to hand. In many cases those who first foresee the opportunity capitalize their foresight and withdraw from active association with the industry, carrying with them legal title to a large share of all future profits. In other cases they remain as active managers in the new enterprise and receive wages of management in addition to profits.

The small scale risk bearers. Our illustrations have consisted in examples of large scale operations of promoters which result in large profit rewards. But we must not allow the glitter of these great prizes to obscure the fact that the service of which we have been speaking and the profit which is its recompense are involved in the most commonplace and picayune of business operations. The

little corner store which appears in the midst of a straggling suburban settlement signifies someone's faith in his foresight; it may be wiped out by a shifting of industries, or a change in transportation; it may flourish on an increasing need for the services of a middleman in the little community. The tailor shop with its small fund of capital, the little factory in which are invested the savings of a group of wage earners, the grist mill on the site of a waterfall, the brokerage offices, dealers in real estate, warehouses, theatres, and hotels are alike affected by the hazards involved in the futurity of income and cost. Every independent enterpriser who embarks on a new line of productive service or adds to the existing facilities in the older lines is to some extent relying on his prevision of future events. Success will bring him excess of income over and above the necessary payments for his labor of management; failure will wipe out this excess of income and may sweep away both the reward for his labor and his entire investment of capital.

Not all risks inevitable. This brief discussion of the source of pure profit is by no means intended as an explanation, much less as a justification, of all income to which the word profit is ordinarily applied. The forecasting of future events and the bearing of the risk involved in them is an indispensable and productive service. But this is not to say that all risk is an inevitable outcome of social forces and that risk-bearing as such is necessarily productive. Risk is often deliberately created so that men may profit by it at the expense of others. The gambler, including the ignorant player of the stock market, certainly takes a risk, but he does not thereby serve society. The man who "rigs" the market or spreads false reports about economic conditions multiplies the risk of other producers and draws an income from their losses. This income he will call profit, but it is clearly in different case from the income of the man whose foresight prepares him to serve the future wants of society or to reduce its future costs. Mere foresight and the willingness to take a chance is not always a social service. The monopolist uses foresight for the purpose of exploiting the necessities of men by contributing to the scarcity of goods and services. The man who preëmpts a choice section of land is exercising foresight, but he has

done nothing to ease our struggle for existence. If we disregard these cases of unproductive income parading under the name of profit and of true profit resulting from anti-social activities, it is not through ignorance of their existence and of the evil they may cause. Our purpose is to emphasize the normal and the essential. We must see that the need for prevision of the future and of risk-bearing is inherent in the very nature of our economic order and that profits are the income resulting from these services.

Profits tend to be short-lived. Pure profit, in each individual case, is a temporary income. The forces of competition work incessantly toward a readjustment of prices which will wipe out the margin of income above costs. Chance gains resulting from lucky ventures or the caprice of fortune are by nature limited in duration to the continuance of the special event which has produced them. Special devices or methods of management, by which individual business men avoid loss from routine business hazards, become known to the trade in general and lead to an adjustment of prices. Even the larger flows of profits resulting from far-reaching changes in industry are of limited duration. A new invention will bestow profits on its first possessor by reducing his costs below these of his competitors. But when it becomes the common property of the trade, all costs will fall and the maladjustment will disappear. The long run effect of improvements in machine equipment and in processes of manufacture is not to expand the profit income but to increase the real incomes of all groups in society. Even the forerunners in new fields of industry, who are the chief profit makers in the business world, are not long exempt from competition. These first comers may withdraw with large amounts of capitalized profits before the levelling tendency of competition has brought prices and costs into adjustment, but the industry as a whole soon ceases to be a source of profit to those who exploit it.

Profits and monopoly. There is some element of monopoly in most cases of profits. The discrepancies in prices which create this income exist because of the lethargy and friction with which the forces of competition operate. Aside from chance gains, the profit-getter prospers because of some competitive advantage which

he enjoys during the interval required for knowledge to spread through the business world or for readjustments to be made in the application of society's labor and capital. Like the monopolist, he draws an income in excess of costs, which a perfect functioning of competition would have denied him. Unlike the monopolist, however, this excess income is not due to his control over the supply of the product.

Profits not a price determinant. Pure profit is not one of the price determining costs of production. This income may not exist as a positive quantity for industry as a whole over a term of years. The service of the profit maker is induced by the chance of receiving the reward, not by the certainty of it. In any hazardous undertaking it is the anticipation of the reward which induces the effort of the competitors. In athletic contests, a winner's prize need be given to but one of the contestants. The Liverpool Sweepstakes enriches one out of a million participants, but the glitter of this single prize is enough to maintain a perennial patronage of this gambling institution. It is probably true that the service for which pure profit is the reward would not be rendered if none of the enterprisers was able to make a profit. But by holding forth chances of gain to individuals, society is supplied with the risk-bearing service of the enterprisers as a class and, if our previous conjecture is true, *i.e.*, that the losses fall little if any short of the profits in the long run, with little or no cost to itself. This class has to bear the losses of business in return for reaping the profits. It should be apparent then that society stands to gain little directly from a social adjustment which would take from business men their profits by relieving them of their risk bearing functions. If profits are appropriated, losses must be borne, and the result might well be a net loss to society.

Receivers of profits control modern industry. We have observed that profits are associated with ownership and not with management. The consequence of largest social importance resulting from this association is that control over business processes lies in the hands of the owners and not in those of the managers and workers. In the typical corporation it is the board of directors

which issues the orders and applies the criteria of success to the operations of the industry. This board is the representative of the stockholders, the owners, and the owners are absentees taking no part in the operations of the business. They do not experience in their own persons the effects of the working conditions which obtain in their industry, and consequently they find it hard to understand and sympathize with the life problems of their employees. Moreover their position is such that they tend to lose sight of the true social purpose of the industry. From their point of view, a business enterprise exists to return a profit and not to render a service to society. As we have seen at an earlier stage in our study, the point of view of the typical laborer is equally self-centred and short-sighted, but the owners are distinguished by the fact that they possess power to give effect to any anti-social policies which may redound to their profit. At all events, since control goes with ownership, a line of division is drawn between the profit-getter and the worker, which places the latter under the direction of the former.

In recent years there has been much criticism of this division of power in the industrial world. Many projects have been launched to bring about a partnership of worker with owner and an equal sharing of power between them. "Industry organized for service" has been advanced as an ideal of greater worth than the present state of "industry organized for profit." These criticisms of the present situation may be based on sound moral principles, and the reforms they suggest may be both desirable and practicable; we shall examine them in detail later. But we must pause here to note that as long as the economic order remains in essentials what it is now the profit-getter must remain the commanding officer of business enterprises.

This follows from the very nature of the risk-bearing service which the profit-getter renders. Considered in its broader social aspects as the function of forecasting future events and guiding the adjustment of society in preparation for them, risk-bearing obviously implies control. Foresight divorced from power can have no practical influence over the adjustment of the industrial struc-

ture. The promoter of new industries must be in a position to give effect to his plans by manipulating the productive factors — land, labor, capital. It is equally true that control must lie in the hand of the bearer of the commonplace risks of routine business operation, if he is correctly to discharge his function. He can defend himself against the evil consequences of risk only if he has the power of command over the enterprise. He can remove or reduce the causes of risk to the permanent gain of all members of society, only if the other productive factors respond to his orders. Until risk disappears from our economic life or until society takes steps to bear the burden of risk coöperatively, the profit-getter must retain control over the industrial structure.

CHAPTER XLI

PERSONAL DISTRIBUTION

Functional shares in distribution. In the preceding chapters we have studied the problem of distribution from the functional point of view. This has involved a classification of the total income fund of society into distinct categories or classes, each category corresponding to some peculiar productive function performed by the recipient of the income or by the wealth which the recipient owns. We may now summarize the result of this analysis. The broadest classification of income from the functional viewpoint discloses two major types of income: income from *ownership* and income from the *labor* or *service* of the recipient. Each of these chief groups contains a number of sub-groups as set forth in the following summary:

FUNCTIONAL SHARES IN DISTRIBUTION

Sharers	*Name of Income*
I. Owners of productive instruments	
A. Land owners	Economic rent
B. Owners of capital	Interest
II. Performers of human service	
A. Laborers (in the broad sense)	Wages (in the broad sense)
1. Employees	Wages (in the narrow sense)
a. Subordinate	Wages
b. Managerial	Wages of management
2. Owner-managers (as workers)	Wages of management
3. Professional workers	Fees
B. Enterprisers (as owners)	Profits

The total money income of the people of the United States. A study of functional distribution will not explain the amount of income received by any chosen individual or social group, since no section of society draws the whole of its income from a single source. An explanation of *personal* distribution involves a study of actual incomes as disclosed by the statistical records which we

possess. Before proceeding to this study however it will be of interest to discover, if we can, how large is the total income at the disposal of society's members.

Most of the attempts to estimate the total money value of the income of the people of the United States have been little better than guesses. The most reliable estimate is that based on a study made by the National Bureau of Economic Research, whose results are brought together in the following table.[1]

TABLE I

NATIONAL INCOME, 1909–1921

Year	*Total Income in Millions*	*Per Capita Income*
1909	27,100	$299
1910	28,400	307
1911	29,000	309
1912	30,600	321
1913	32,000	329
1914	31,600	320
1915	32,700	326
1916	39,200	385
1917	48,500	470
1918	56,000	537
1919	67,254	640
1920	74,158	697
1921	62,736	579

Care must be taken when interpreting the meaning of the figures in the "per capita" column. Of course these figures are no indication of the average earnings of the workers or the average income of any other social class; the total income of the nation is averaged over the total population, including many who received no individual incomes: non-workers, infants, the aged, the crippled, etc.

The table shows a startling increase in total income during the twelve year period, 1909–1920, the total rising from a little more than 27 billions, to more than 74 billions of dollars. The per capita income, because of the increase of population, rose with

[1] *Income in the United States, 1909–1919*, Vol. I, p. 143, and *Income in the Various States, 1919, 1920, and 1921*, p. 32. The figures in the first of these volumes were later revised by the National Bureau of Economic Research and the revised estimates issued in 1925. Our tables are based on the revised estimates.

slightly less rapidity, from $299 to $697 per capita. The decline in both these figures between 1920 and 1921 is accounted for by the business crisis and the ensuing depression which appeared late in 1920.

Total income corrected for price changes. However the increase in total and average welfare was more apparent than real. When income is stated in terms of money value the figures are affected by changes in the general level of prices. The period embraced in the table was one of rapid inflation of prices, reaching its peak in 1920. The trend of real income over a period of years is shown more accurately when money values in the different years are reduced to a common standard of purchasing power. In the following table this is done by translating the figures into terms of the price level of 1913.[1]

TABLE II

THE NATIONAL INCOME, 1909–1921
STATED IN TERMS OF 1913 PRICE LEVEL

Year	*Total Income in Billions*	*Per Capita Income*
1909	28.2	$312
1910	29.1	315
1911	29.3	312
1912	30.8	323
1913	32.0	329
1914	31.3	316
1915	32.0	319
1916	35.5	349
1917	37.3	361
1918	35.5	340
1919	37.6	358
1920	36.3	341
1921	36.2	334

Comparative national incomes. It would be interesting to know the comparative status of the important industrial nations in respect of the average economic welfare of their peoples. This is a subject on which we have little reliable information and that consisting in statistics which make comparisons between the different countries a difficult undertaking. The following table

[1] *Op. cit.* See previous footnote.

gives the figures for total and per capita money incomes of certain modern nations in 1914, assembled from estimates made by different statisticians and at different times.[1]

TABLE III

ESTIMATED TOTAL AND PER CAPITA NATIONAL INCOMES OF VARIOUS COUNTRIES, 1914

Country	*Total Income in Billions*	*Per Capita Income*
United States	33.2	$335
United Kingdom	10.9	243
Germany	10.5	146
France	7.3	185

The table shows that the people of the United States enjoy a much larger per capita income than those of any other great industrial nation. This is undoubtedly true to the facts, since the abundant national wealth of this country, together with the relative scarcity of the population, result in a large average product from the economic effort of our people. But the figures can be used only as rough indications of the economic welfare of these different national groups. They have not been corrected with reference to the different price levels and hence do not indicate comparative real income. Still less do they prove a higher degree of contentment and happiness in the countries having the largest average money incomes. Contentment and happiness are psychological magnitudes which are as much affected by the range of wants and the customary mode of life which we call "the standard of living" as they are by the absolute incomes of the different peoples. Moreover our figures for the average or per capita income are obviously abstractions, which do not necessarily indicate the economic status of any section of the population. The actual incomes of the people depend upon the type of personal distribution which obtains in the country. A country with a relatively small average income distributed with approximate equality would have a larger section of its people living on a plane of relative comfort than would a country with a larger average income most of which was received by a small upper class.

[1] Taken from *Income in the United States, 1909–1919*, Vol. 1, p. 185.

Inequality of personal incomes in the United States. All studies that have been made of actual income distribution in modern countries show striking agreement in one regard; they all indicate a condition of extreme inequality. Even in the United States, where economic opportunity is relatively equal as compared with older countries, there is a much wider range of variation between the incomes of the small class of well-to-do at the top of the scale and the mass of the people at the bottom than might be supposed by one who had not made a careful study of the facts. How wide is this range of variation in personal incomes is shown by Table IV. This table, based primarily on information contained in income tax returns, consists of an analysis of incomes actually received by individuals in the United States. It should be noted that these figures have no relation to the per capita averages listed in Table I.

TABLE IV

DISTRIBUTION OF MONEY INCOMES IN THE UNITED STATES IN THE YEAR 1918[1]

Annual Income Less Than	*Cumulative Percentage of Persons*	*Cumulative Percentage of Incomes*
$ 600	9.5	2.4
800	22.4	8.3
1,000	38.7	17.9
1,200	54.5	29.1
1,400	67.1	39.7
1,600	76.0	48.3
1,800	82.0	54.9
2,000	85.9	59.7
2,500	91.4	67.5
3,000	94.1	72.3
4,000	96.6	77.9
5,000	97.8	81.2
6,000	98.4	83.5
8,000	99.0	86.3
10,000	99.3	88.0
15,000	99.6	90.5
25,000	99.9	93.6
50,000	99.94	95.3
100,000	99.98	96.9
500,000	99.9996	99.1
1,000,000	.0004	99.5
Over 1,000,000		.5

[1] Adapted National Bureau of Economic Research, *Income in the United States, 1909–1919*, p. 134.

In 1918 approximately 10 per cent of the income receivers had annual incomes of less than $600, and this group as a whole obtained 2.4 per cent of the total national income. About 86 per cent of the people had incomes of less than $2,000 and received as a group about 60 per cent of the total, and so on through the table. It should be noted that these figures refer to the year 1918, when economic conditions were much distorted by the war. This factor however probably had the effect of reducing the inequality of distribution rather than of exaggerating it. Comparison with similar studies of personal distribution for other years shows that the war period was one in which the lower incomes rose faster than the higher. In normal times therefore the proportion of the total income of the United States which accrues to the small upper class is even larger than is shown by the table.

The condition of inequality pictured in these figures may be thrown into clearer light if we group the individual income receivers into general social classes. If we begin with the highest incomes and mark off the upper one per cent of the people, we find that this small group received 14 per cent of the total income in 1918. Going back to the more detailed figures of which Table IV is an adaptation, we may isolate the upper 5 per cent of the people and find that they received 26 per cent of the total income in this year. This method of presenting the condition of unequal distribution, if continued to include all income receivers, will result in the following table.

TABLE V

DISTRIBUTION OF THE NATIONAL INCOME 1918
BY SECTIONS OF THE POPULATION

Section of the Population	*Percentage of Total Income*	*Range of Actual Income*
Richest 1%	14%	$8,000 to more than $1,000,000
Richest 5%	26%	$3,300 and more
Richest 10%	35%	$2,400 and more
Richest 20%	47%	$1,800 and more
Poorest 39%	18%	Below $1,000

If we draw a line approximately midway through the national income, we find that about one half (47 per cent in 1918) accrued to one fifth of the people; while the remaining half (53 per cent)

was shared among the other four fifths. The 39 per cent of the people having the lowest incomes received only 18 per cent of the total national income. The third column of the table is noteworthy since it throws light on the modest size of the money income received by the mass of people in this country. In order to mark off the upper one per cent of the income receivers it is necessary to include incomes as low as $8,000; the richest 5 per cent includes those whose incomes were as low as $3,300; and so on through the table. Almost 40 per cent of the people received incomes of less than $1,000.

Inequality universal. Similar conditions of inequality are to be found in all countries for which studies of personal distribution have been made. It is not our purpose to assemble all the statistical data on this subject, nor to draw a detailed picture of income distribution in all the different countries. As evidence of the general type of distribution on the continent of Europe, the statistics of income taxation in Prussia are illuminating. Taking the figures for the year 1908, we find that out of a total population of about 38,000,000, 45 per cent were exempt from the income tax because their incomes were less than 900 marks ($225 per year). Nine tenths of the taxpayers had incomes of less than 3,000 marks; only one tenth, incomes exceeding this amount. A similar analysis of the British income tax returns for 1904 shows that less than 10 per cent of the people received half of the total income of the country. These figures are typical of all modern countries. Everywhere the largest number of individual incomes is massed at the lower end of the income scale; while the small number at the upper end of the scale receive the lion's share of the total income.

Perhaps it is well to point out that these figures for the distribution of money incomes do not indicate the comparative scales of consumption or enjoyment of the different individuals and social classes. The upper middle and the highest income classes consistently save and invest a portion of their money incomes; the lower income classes save very little. Saving reduces the amount of current consumable income, and continued reinvestment of the

income from savings leaves the recipient in the same position, in respect of consumable income, as if his original money income had been proportionately smaller. In the case of many of the largest incomes, this process of continuous postponement of consumption is carried on throughout the lifetime of the income receiver. To what extent the practice of saving affects personal distribution of final income we do not know, since we have no accurate statistics of the amounts saved by different social groups, but there can be no doubt that this factor has the effect of modifying, in some degree, the condition of extreme inequality pictured in our statistics of money incomes. It tends however to accentuate inequality in the distribution of wealth (as distinct from income), while, coupled with the institution of inheritance, its ultimate effect is to perpetuate inequality of both income and wealth through succeeding generations.

Inequality graphically presented. The scheme of income distribution which has thus far been presented in the form of statistics may be thrown into clearer light by the use of a simple graph. In the following diagram, for purposes of comparison, two curves are drawn on the same scale: the straight diagonal line, representing a hypothetical condition of uniform income distribution, and the curved line plotted from the income statistics for the United States in 1918.

Examination of the " line of equal distribution " will show that it represents a condition in which 10 per cent of the income receivers get 10 per cent of the total income; 20 per cent of the income receivers, 20 per cent of the total, etc. The sagging of the curve for 1918 income below this line indicates the existing inequality in distribution. To illustrate, following the perpendicular from the point on the base axis which represents the poorest 50 per cent of the income receivers, it will be seen that it cuts the curve at a point which indicates approximately 26 per cent of the total income (indicated on the horizontal axis). Similarly, the poorest 60 per cent of the income receivers got about 34 per cent of the total income; 70 per cent of the income receivers, 43 per cent of the total, etc. The general shape and trend of this curve is the same for all

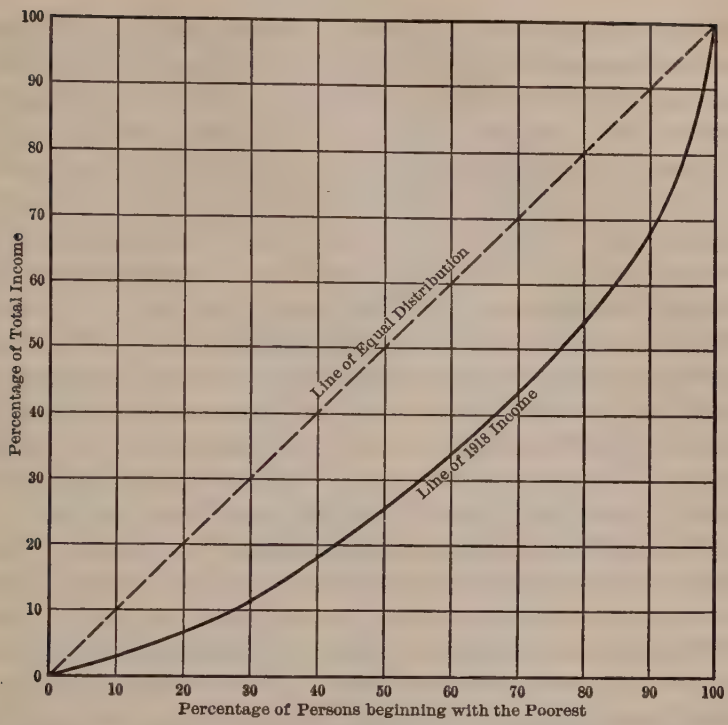

FIG. 7. — Curve of Income Distribution in the United States, 1918.[1]

countries, though its slope may differ to some extent to conform to local conditions. It may be well to caution the reader that the straight diagonal line which represents uniform distribution is not intended to picture an ideal condition of affairs. The question of what constitutes a perfect type of distribution is one which cannot be decided on economic grounds alone; it raises problems of a moral and ethical nature which lie outside our field. It is probably true however that none of our accepted precepts of justice would prescribe the condition of flat uniformity pictured in the straight line on our graph.

Personal distribution of wealth. Thus far, we have considered the distribution of income, a problem quite different from that of the distribution of wealth. Our statistics showing the division of

[1] Taken from *Income in the United States, 1909–1919*, Vol. I, p. 141.

property rights among the people of modern nations are meagre and fragmentary, but such information as we possess points to the conclusion that the wealth of the world is even more unequally divided than is income. This is not unnatural, considering the fact that a large proportion of the social income is divided among the wage-earning classes which on the whole possess no property. The division of income is a much more accurate index to the economic welfare of individuals and social groups than is the division of wealth, since it is from income alone that the wants of mankind are satisfied. Nevertheless a knowledge of the state of wealth distribution is not without interest. In a society sanctioning private property and relying upon individual enterprise for economic services, control over the industrial structure lies in the hands of the property owners. A wide diffusion of property rights in such a social order makes for stability and internal order.

Evidence from American probate courts. The records of the probate courts, through which estates pass on the death of their owners, give some indication of the distribution of wealth among the people of the country. It cannot be assumed of course that such records are accurate in detail. Family estates are often parcelled out before the death of the head of the family. The average age of those who die in a given year is higher than the average for the living population, and the decedents in consequence have had a longer time in which to accumulate wealth. There are other factors which must be taken into account when an attempt is made to use the statistics of the probate courts for purposes of estimating the actual distribution of ownership. Despite these qualifications however these statistics may be accepted as indicating in a broad way how the people of the country share in its store of accumulated wealth.

In 1915 a careful study was published of the records of the probate courts of Massachusetts covering a period of sixty years from 1829 to 1891. The same study contained an analysis of the wills probated in certain sections of Wisconsin in the year 1900. It shows a marked similarity in the evidence which these two groups of statistics afford regarding the distribution of wealth in the

United States. This evidence is presented in a simplified form in Table VI.

TABLE VI

SECTIONS OF THE POPULATION CLASSIFIED ACCORDING TO WEALTH OWNERSHIP [1]

Section of Population	*State*	*Per Cent of Total Wealth*	*Average Value of Wealth*
Poorest 65%	Mass. 1891	4.5	$399
	Wis. 1900	5.2	381
Lower Middle 65–80%	Mass. 1891	3.9	1,499
	Wis. 1900	4.8	1,524
Upper Middle 80–98%	Mass. 1891	32.8	10,509
	Wis. 1900	33.0	8,730
Richest Class 2%	Mass. 1891	58.8	169,550
	Wis. 1900	57.0	135,715

The poorest two thirds of the people owning little more than a few hundred dollars worth of furniture and personal effects each, had a combined ownership of 4 or 5 per cent of the total wealth. The next 15 per cent had an even smaller share of the total. These two groups together, representing 80 per cent of the people, owned less than 10 per cent of the total wealth. The remaining fifth of the people owned 90 per cent of all the wealth, and of this group a thin upper layer — two per cent of the total population — owned almost three fifths.

Personal distribution of wealth in other countries. The same authority has collected the available statistics of wealth distribution in France, the United Kingdom, and Prussia. The information for the United Kingdom and France is based on probated estates; for Prussia, the statistics are taken from tax returns. The estimated distribution of wealth in these three countries is shown in a simplified form in Table VII.[2]

[1] W. I. King, *Wealth and Income of the People of the United States*, p. 79.
[2] King, *op. cit.*, p. 96.

TABLE VII

SECTIONS OF THE POPULATION OF PRUSSIA, FRANCE, AND THE UNITED KINGDOM CLASSIFIED ACCORDING TO OWNERSHIP OF WEALTH

Section of Population	*Country*	*Per cent of Total Wealth*	*Average Value of Wealth*
Poorest 65%	Prussia 1908	4.9	$153
	France 1909	4.3	186
	U. K. 1908	1.7	133
Lower Middle 65–80%	Prussia 1908	5.5	743
	France 1909	5.6	1,052
	U. K. 1908	2.9	979
Upper Middle 80–98%	Prussia 1908	30.6	3,445
	France 1909	29.4	4,602
	U. K. 1908	23.7	6,670
Richest Class 2%	Prussia 1908	59.0	59,779
	France 1909	60.7	85,500
	U. K. 1908	71.7	181,610

There are certain minor differences in detail shown by these figures from the different countries, but the similarities are much more striking. Everywhere the general scheme of wealth distribution is the same; a minute upper fraction of the people owning considerably more than half of the total wealth, a large major fraction of the total population at the bottom of the scale owning virtually nothing, between these two extremes a lower middle class with a fraction of the total wealth much less than proportionate to its numbers and an upper middle class with somewhat more than its proportionate share. All other evidence which we possess points to the same general conclusion regarding wealth distribution in nations of advanced civilization.

Labor's share of the national income. In view of the extreme inequality which exists in the distribution of wealth, it is of some importance to know what fraction of the total income accrues to the wealth-owning classes, and what fraction is distributed as a reward for labor. This is a subject on which we have no precise information at all. Individual incomes are often drawn from

several different sources and represent several types of productive function. It is impossible to classify the income receivers of the country under the four headings which we have adopted for our study of functional distribution — landowners, capitalists, laborers, enterprisers. Almost the only functional distinction which we can draw between the income receivers of the industrial world is that which separates the hired workers on the one hand from all other sharers in the product of industry on the other.

In the following table, this division is made between the wages of hired workers and other types of income on the basis of statistics gathered by the National Bureau of Economic Research.[1] The figures apply to the year 1913, the last normal year preceding the war. It should be pointed out that the percentages do not measure the total wage share of the social income; they refer to explicit wages only, leaving out of account owners' wages of management and the implicit wage income of self-employed workers.

TABLE VIII

PERCENTAGE OF INCOME OF VARIOUS INDUSTRIES DISTRIBUTED AS WAGES TO HIRED WORKERS IN 1913

Agriculture	13.4%
Mining	73.4
Factories	74.5
Hand trades	66.7
Railways	66.4
Public utilities other than railways	52.9
Water transportation	79.1
Banking	31.6
Government	91.7
Miscellaneous	63.2
Average	55.6

Something over half of the total product is paid to hired workers in these industries as a whole. But this average is kept down by the abnormally small percentage of wages, in proportion to other shares, in agriculture, where most of the work is done by the farmer and the wage income is implicit, and in banking, where the proportion of hired labor to capital is small. In those branches of

[1] *Income in the United States, 1909–1919*, Vol. I, p. 92.

industries which may be taken as typical of a highly industrialized society — manufacturing, mining, and transportation — the share allocated to the wage-earner varies between two thirds and three quarters of the total product. The remaining fraction of the income includes interest, rent, profits, and implicit wages of management.

The problem of poverty. Despite the immense increase in the physical productivity of industry during recent history, poverty has not disappeared. Indeed even in so prosperous a country as the United States the majority of the people live on an economic plane which, tested by our present standards of wholesome living, represents a condition bordering on poverty. Outright pauperism and destitution is the lot of only a minute fraction of the people and is recognized by organized society as an intolerable evil to be alleviated by public and private relief. But these cases of individual misery and suffering are of trifling importance as compared with the vast number of self-supporting and industrious people who find it impossible to maintain what is now known as a wholesome mode of life on the basis of their share in the social income. If we refer again to Table IV, which shows personal distribution of income in the United States, we note the fact that 39 per cent of all the incomes in 1918 were less than $1,000. If a line is drawn through the $2,000 income, it will be found that 86 per cent of those gainfully employed received less than this amount. These are small incomes when judged by our present standards of living and the present level of living costs.

Family incomes and living costs. A better indication of the welfare of the working population is furnished by the statistics of family incomes in comparison with living costs. The family is the economic unit in society. The normal family is assumed to consist of five members, two adults and three dependent children, since an average of three children per family is required to maintain the level of the population. It is further assumed that this hypothetical normal family is supported out of the income of its chief wage-earner and has no income from sources other than the sale of labor. These assumptions are not entirely true to the facts.

The average wage-earning family appears to have an income from property which amounts to between five and eight per cent of its total income, though these supplementary receipts vary directly with the wage income, being of least importance in the lowest wage groups. The other assumption — that the family is supported on the income of its chief wage-earner — is still farther from the truth; child and woman labor are customary on the lower ranks of the laboring population. These qualifications must be borne in mind when attempting to estimate the economic status of our people by comparing the actual wage incomes of laboring families with their costs of living.

Minimum standard of living in America. There have been many careful estimates of the cost of a minimum standard of wholesome living for a family of normal size. In 1914 Professor Hollander stated: "That in order to maintain a decent standard of living in the United States for an average family of five, an annual income of $600 to $700 is insufficient; that $700 to $800 requires exceptional management and escape from extraordinary disbursements consequent upon illness or death; and that $825 permits the maintenance of a fairly proper standard." At the same time the Board of Estimate and Apportionment of New York City estimated that $845 was the minimum for unskilled laborers in that city. According to the New York Factory Investigating Committee, $876 was the minimum in 1914 for the State as a whole. These estimates, though differing slightly from each other, give a sufficiently accurate idea of the cost of living for wage-earning families in the last year preceding the war.

American family incomes. At this time the actual annual earnings of industrial workers in Pennsylvania, as reported by the Department of Labor and Industry of that State, averaged $720. The Industrial Commission of New York estimated that the existing wages for factory and office workers would return an annual income of $649 if the workers suffered no unemployment. In 1915, Lauck and Sidenstricker carried out an exhaustive analysis of the actual incomes of wage-earners over the country, which resulted in their conclusion that four fifths of the family

heads in the working population of the United States received less than $800 per year.[1]

The United States Bureau of Labor Statistics has estimated that the cost of living rose by 112 per cent from 1914 to 1920 for the country as a whole. Throughout all this period the advance of wage rates lagged behind the rising price level. In 1920 an income of approximately $2,100 was required to maintain the estimated standard of living for a normal family of five. In this year the full time earnings of factory workers in New York State average less than $1,500, according to the findings of the State Industrial Commission. The average annual wage of unskilled labor on the railroads, which may be taken as an index to the general economic condition of this lowest labor group, was approximately $1,250 in 1920. Even in the strongly unionized trades average wages rose less than 100 per cent on the average as against a rise of 112 per cent in the cost of living during this period. By the end of 1923 the business of the country had reached a stable position after the violent fluctuations of the preceding decade. At this time the cost of living was 73 per cent higher than in 1914, thus requiring an income of approximately $1,400 a year, or about $27 per week to maintain the standard of living for a wage-earning family. At this time, according to findings of the National Industrial Conference Board, the average wage per week for the country as a whole was $25.12; for skilled male workers, the average was $28.40; for unskilled, $21.31.

Qualifications and conclusions. Here we have a brief history of the real income of a normal wage-earning family in this country during a decade of rapid economic change. It is not asserted that these figures do more than indicate, in a general way, the relative poverty of the mass of our laboring population as measured against a plane of living acceptable under the high standards of our advanced civilization. A crude average of widely varying wage

[1] A collection and analysis of the different available studies of earnings and living costs for American wage-earning families is given in an article by Abraham Epstein, "American Wages, 1890–1920, and the Standard of Living," *Annals of the American Academy of Social and Political Science*, Sept. 1921. The figures quoted are taken from this article.

rates is no index to the real economic position of any particular family or labor group. Moreover, actual earnings are not correctly reflected in the wage rate; account must be taken on the one hand of unemployment, accident, sickness, and other misfortunes which reduce earnings below the amount indicated in the wage rate, and on the other hand of overtime pay, bonuses, and similar receipts which increase earnings. Again the assumption of a normal family of five is not true to the facts. Investigations shows that in 1920 about 27 per cent of all adult males gainfully employed were unmarried, and it is estimated that two thirds of the wage-earning families consist of less than five members. Finally, a great many families in the wage-earning population are not supported by a single worker. In 1920, 6,000,000 adult women were gainfully employed in the United States and 6,300,000 youths between the ages of 16 and 20 years.

Making all due allowance for these qualifying factors, however, it still remains true that the family incomes of laborers in the United States, notably in the lowest groups of unskilled and semi-skilled workers and to a less degree in the group of the "white collar" workers, provides very little margin above the poverty line of existence.

Problems raised by inequality. It is not necessary to attempt the impossible task of formulating rules to determine the "best" type of income distribution in order to see that extreme inequality is an evil. The only measure we have of the final result of society's economic effort runs in terms of utilities derived by the ultimate consumers. We may safely infer from the general law of diminishing utility that extreme inequality of distribution reduces the aggregate of human satisfactions derived from a given amount of social income. This does not prove that flat uniformity of distribution is an ideal state of affairs from the standpoint of society's long run interest, since this would be to reward men without regard to the economic value of their services and hence to discourage initiative and industry. But it suggests that any lessening of extreme inequality, resulting from a free play of economic forces, will increase the benefit from society's productive effort.

Furthermore there can be no doubt that the economic welfare of a social group over a period of time is vitally affected by the stability of its institutions and the smoothness with which they function. This is determined in large degree by the contentment of the average man with his lot in the social structure. Discontent and rebellion are engendered, not so much by absolute poverty and misery, as by inequality of fortune, the relative poverty of the majority in contrast with the opulence of the few. It is now generally admitted that any arrangements which tend to accentuate inequalities between social classes are impermanent. They develop resentment and resistance which will ultimately destroy them. And the process of the modification of such social institutions is likely to be accompanied by such turmoil and disorder as will tend to impoverish the whole social group.

Is inequality increasing? For these reasons it is important that we discover, if we can, how distribution on former stages of social development compares with that of today and whether the trend in our capitalistic society is in the direction of greater inequality or the reverse.

Many of the statements which purport to throw light on this question are beside the point. For example it must be admitted that the level of comfort for all classes in modern industrialized countries is higher than at any other period in the history of these countries and higher also than the level of comfort in modern non-industrialized nations. The poor in a capitalist society enjoy satisfactions which were formerly within the reach only of the wealthy. Absolute misery and destitution are on the decrease and have been decreasing throughout the whole of modern history, if we except those periods in which civil anarchy or international war prevailed. But this does not prove that inequality is diminishing.

Growing inequality in America. In the United States inequality is probably greater now than it was a century ago. Like all new countries, we began our life with an abundance of natural resources and a relative scarcity of labor and capital. The share of the social income allocated to ownership was small. Most incomes were drawn from labor, and most laborers were given opportunity to

apply their effort to natural resources which were practically unlimited. It is the competition of life, the struggle of man with man for control over limited income-producing opportunities, which brings out differences in the inherent abilities of individuals and produces inequality of fortune. On the frontier most people lived in a condition of rude comfort on a plane of comparative equality. The passing of the frontier, the multiplication of the population, and the growth of capitalism destroyed this condition.

During the nineteenth century all classes increased their prosperity, but the rate of increase for the well-to-do was faster than the rate for the poor. The appearance of capitalism in industry, with its immediate tendency toward large scale production, had the effect of concentrating income in the hands of the possessing classes. The first comers in the field of business organization reaped large rewards in the form of profits. Rising rents and land values contributed to the income from ownership, especially in urban enterprises and in certain sections of the mining industry. At the same time floods of immigrant labor retarded the rise of wages for the mass of workers at the bottom of the scale. The outcome of these forces during the last half of the nineteenth century is to be seen in the rapid multiplication of large fortunes. For the lower and middle classes economic improvement was steady, though less rapid. It would be unsafe to conclude however that this early tendency toward increasing inequality will continue now that the country has attained economic maturity. For evidence as to the normal effect of growing capitalism on the relative prosperity of different social classes, we must turn to the record of older countries.

Evidence from other countries. As in other phases of the problem of personal distribution, our information on this question is disappointingly meagre. The classic analysis of the trend of distribution over a recent period was made by A. L. Bowley in *Change in the Distribution of the National Income, 1880–1913,* published in 1920. This study is based on the income tax returns of Great Britain. Its net result is that no significant change has occurred in the status of distribution during the period under

review. All classes shared about equally in the increasing prosperity of the country; the well-to-do class received about the same share in 1913 as in 1880; the fraction going to property, as compared with that accruing to labor, was also unchanged.[1]

Examination of the tax returns of Germany during the same period shows substantially the same result. The average income rose, and all classes shared in the increase. In this country, as in the United States, because of the more rapid pace of industrialization there was greater tendency toward increased inequality than in England. The incomes of the wealthy classes of Germany increased at a relatively faster rate than did those of the poorer classes, but the difference was not sufficiently striking to support a conclusion that increasing inequality is an inevitable characteristic of our system of industry. What other information we possess agrees with the results of these studies.[2] Karl Marx's charge against capitalism, that it increases the riches of the wealthy and the poverty of the poor, though it is not definitely overthrown by our knowledge of present tendencies, must at least be dismissed as "not proved."

Causes of inequality. It is customary to divide the causes of inequality into two groups: (*a*) those which inhere in the character of men, and (*b*) those that arise from the nature of the institutions of society. Writers on this subject differ in the amount of stress which they place on these two factors. There are those who trace all differences in material fortune to differences in the natural endowments of individual men. This conviction leads to two conclusions, both in defense of the present order of things: first that existing inequalities are a fairly precise measure of differences in the economic services rendered by different individuals and social groups, secondly that the present state of distribution is determined by natural law and cannot be materially altered.

On the other hand there are many who believe that existing in-

[1] A. L. Bowley, *op. cit.*, p. 22.

[2] Pareto, *Cours d'économie politique*, Vol. I, contains studies of the distribution of income in different countries chiefly in the nineteenth century. These studies show a marked similarity in the scheme of distribution in different places and at different times.

equalities are created and perpetuated by social arrangements. The rich owe their favored position to privilege sanction by social force rather than to their own intrinsic merit. Those who hold this view are inclined toward reform, believing that there is no obstacle other than established custom in the way of intentional redistribution of income in the direction of greater uniformity.

We may note in passing that this attempt to distinguish between social and individual responsibility for inequality is not logically sound. In the final analysis all causes of inequality have their roots in social institutions. This is not to deny the existence of those differences in natural abilities which are exalted as the chief causes of inequality of income by the first group of writers to which we have alluded. Undoubtedly men are born with diverse natural endowments. They differ in physical strength and endurance, in native intelligence and ability to acquire knowledge, and, of chief importance from the standpoint of this problem, in perseverance, acquisitiveness, and ruthlessness in pursuit of material gain. But these admitted differences in individual characteristics would not of themselves and inevitably lead to differences in income. They would do so only in a society whose institutions included a system of free enterprise and private property. Authoritative distribution might establish a condition of equality despite individual differences such as we have mentioned, or a condition of inequality which had no references to these differences.

Personal ability and inequality. In our social order as it now is difference in personal ability is one cause of inequality of income, but it is not the only one nor necessarily the most important. On this point we cannot do better than quote the language of Professor Pigou.

"The essential characteristic of current income distribution is that the great bulk of incomes are massed together near the lower end of the income scale. This fact is significant for the following reason. There is clear evidence that the physical characters of human beings — and considerable evidence that the mental characters — are distributed on an entirely different plan. When, for

instance, a curve is plotted out for the heights of any large group of men, the resulting picture will not be, as with incomes, more or less a hyperbola, but it will be a symmetrical curve shaped like a cocked hat. It will . . . be symmetrical about the mean, in such wise that there is no massing near either end, but equal numbers of heights above the average and below it, and a lessening number of people at every height as the distance from the average in either direction is increased. Now, on the face of things, we should expect that, if, as there is reason to think, people's capacities are distributed on a plan of this kind, their incomes will be distributed in the same way. But the fact is that they are not distributed in the same way; and a little reflection shows at once why they are not. The reason is that income depends, not on capacity alone, but on a combination of capacity and inherited property, and inherited property is not distributed in proportion to capacity, but is concentrated upon a small number of people not selected in accordance with their own, or even, in many families, with their parents' capacity, but owing to their good fortune, perhaps to their being only sons or daughters, perhaps to some other 'accident.' The significance of this problem from the standpoint of our present problem is obvious. If the form of income distribution is partly determined by facts of bequest and inheritance, the particular form which is found to be dominant in current conditions cannot possibly be *necessary*, except upon the assumption that the general scheme of inheritance now generally in vogue is maintained." [1]

Inheritance and inequality. No social order whose economic structure is founded on individual initiative and free enterprise can avoid inequality of personal incomes, as long as men begin the competition of life unequally endowed by nature. But in our society the actual scheme of distribution is distorted by the influence of a factor which operates capriciously and without reference to natural abilities. The institution of inheritance acts in each generation as an independent cause of inequality and, through a period of time, as a cumulative force perpetuating past differences in income. Any attempt to rectify in a thorough

[1] A. C. Pigou, *Economics of Welfare*, Macmillan, 1920, p. 696.

manner the existing state of distribution will be compelled to face the question of the merits of these fundamental institutions of society.

Reform of inequality. With the spread of the ideals of democracy through the modern world, all social arrangements which result in inequality of economic status and of opportunity have been put on the defensive. The revolt against the scheme of distribution has gathered such strength that all modern industrial nations have begun to experiment with various programs of reform. We shall not attempt in this place to do more than summarize the general characteristics of these different programs.

Charitable relief. In a group by themselves we may place those devices which are intended, not so much to remedy inequality, as to relieve the distresses of the poorest members of society. Official and private charitable relief was never so abundant in the history of the world as at the present time. This is not due entirely to an elevation of human character; in large part it is a reflection of the increasing average prosperity of the people, which throws each case of misery and destitution into clear relief. It is doubtful whether charity alone will cause any permanent improvement in the problem of poverty, and clear it is that indiscriminate charity must be reckoned among the causes of that problem. Much progress has been made in the administration of charity, both public and private, toward the end that relief of poverty may have a constructive effect in building up the self-dependence of destitute families and individuals. But society has not yet recognized what seems to be a necessary implication of a condition of dependency; namely, that the individual who is unwilling or unable to bear the cost of his own maintenance has no right to marriage or to the procreation of children after marriage. The movement toward the prevention of marriage among the hopelessly unfit and toward the control of births in destitute families has scarcely grown beyond the stage of academic discussion.

Relief from the hazards of life. Similar in nature to the relief of distress, but of more importance in that they attack the causes and not the symptoms of this evil, are those devices which aim to

forestall or eradicate the hazards of life which are so frequently the immediate cause of poverty. Sickness, unemployment, accident, and premature death are normal risks of the modern workingman's life. In the lowest income classes, where the normal wage provides no margin above necessary living costs, the occurrence of one of these disasters is often sufficient to destroy the economic basis of the family. Most of these hazards are insurable risks. When the individual subject to them cannot or will not bear the cost of insurance, it is in society's interest either that he be compelled to do so or that insurance be provided for him at public expense. Compulsory social insurance has so completely demonstrated its utility as a partial solution of the problem of poverty that it has been accepted as a normal feature of industrial life in all modern civilized nations.

Preventive medicine, public clinics, public sanitation, and health movements may be considered of the same social significance as compulsory social insurance. They are safeguards, built up at public expense and in the public interest, against the hazards of life which threaten the lowest social classes with misery and destitution.

Public education. Aside from these devices for the relief or prevention of outright poverty, the programs which aim at greater equality in distribution must work some more or less far reaching change in the existing institutions of society. The general program for free, compulsory education is perhaps an exception to this statement. To the degree that education breaks down the barriers between labor groups, it will tend to lower the higher wage income and raise the lower, at the same time increasing the average. But we do not know to what extent universal education will have this effect of promoting transference from one labor group to another, since it has never been demonstrated that the differences in productive ability which now divide the working population into classes are not due to inborn characteristics. In any event education will not remove inequality in distribution since it will not equalize the capacities of individuals and will leave untouched the institution of inheritance, which at present prevents an equal-

ization of opportunity. Thoroughgoing change in the present scheme of distribution involves a modification of our existing institutions of personal liberty and private property.

Limitations upon individual freedom. We will reserve for later discussion the program of the socialists, which contemplates a revolution in both of these institutions. It should be noted however that all capitalistic countries have begun to modify the operation of these institutions with a view toward bringing about a greater diffusion of income and wealth. Individual liberty and free enterprise are no longer uncontrolled and unlimited. Many business practices which were permissible a generation ago are now forbidden because they enrich the strong or the unscrupulous at the expense of other members of society. Governmental bodies have taken over the ownership and operation of certain forms of enterprise which were formerly a source of private profit. The powers of the monopolist are more and more circumscribed by public regulation of his prices. The weakest bargainers in the laboring classes are protected in many countries by minimum wage laws which prevent their exploitation. This many-sided movement toward restriction of individual freedom in economic affairs is tending to prevent the enrichment of individuals by methods which are condemned by prevailing standards of justice and fair play.

Restriction of property rights. Similarly property rights are being modified by the movement toward greater equality of income and opportunity. Graduated income taxes, in conjunction with the expansion of governmental expenditure along lines of general public service, result in an indirect transference of income from the rich to the poor. The whole institution of inheritance is under strong attack. At present conservative governments have not gone beyond the point of subjecting inheritances to graduated taxes which tend somewhat to reduce the largest fortunes. These devices for a forcible redistribution of existing incomes and existing accumulations of wealth contain elements of danger to society. If pressed too far they will tend to discourage industry among the most capable individuals and to penalize the thrift which leads to

capital accumulation. The same cannot be said of the movement to appropriate the unearned increment of land — in the past a cause of many large fortunes — for public benefit through taxes on rising rents and land values. As we discovered in another place, this device, if used with caution, will leave unchanged present incentives to industry and thrift and will destroy private appropriation of an income which is created by social forces.

All of these methods of modifying the operation of our industrial system in the direction of greater equality of income and opportunity may be illustrated from the practices of any modern government. Whether or no they will be carried to the point of revolutionizing the framework of the capitalistic order is a question whose answer will probably be revealed by the developments of the next succeeding decades.

CHAPTER XLII

POPULATION

The problem of population. Population presents a double problem: on the one hand is the question of number and increase, on the other is the question of capacity and social worth of the individuals composing the group. The quantitative aspect was the first to attract attention and has been the more clearly apprehended and the more widely discussed. It may present a practical problem either because of a paucity or because of an excess of numbers. It has a definite relation to labor problems in that the supply of labor depends on the increase in the numbers of mankind. The second or qualitative side of the problem has, until recently, aroused interest only sporadically. Not increase in numbers, nor national power and glory, but a higher level of popular capacity and welfare, is the quality element of the problem.

Malthusian theory of population. The first writer to set out a theory of population in detail and to support it with evidence was Thomas Robert Malthus, who published in 1798 his *Essay on the Principle of Population.* He was the first to arouse public interest in the subject. Indeed his book made an impression that few books have ever made. Since his time every educated man has had the problem of numbers in mind, and current opinion still reflects, if not in detail at least in general, the view put forward by Malthus.

According to his own account, Malthus attempted to show three things: (1) population is necessarily limited by the means of subsistence, (2) population invariably increases where the means of subsistence increase, unless prevented by some very powerful and obvious checks, and (3) these checks which repress the superior power of population and keep its effects on a level with the means of subsistence are all resolvable into moral restraint, vice, and misery.

The whole basis of the Malthusian argument is the comparative

rapidity of population increase and of the increase of food. Population, Malthus said, when unchecked increases in a geometrical ratio, doubling in approximately twenty-five years; food increases in an arithmetical ratio. For example, population would increase in two centuries as the numbers 1, 2, 4, 8, 16, 32, 64, 128, while the supporting power of the soil was increasing as the numbers 1, 2, 3, 4, 5, 6, 7, 8. Hence there would be a growing intensity of pressure of population on the means of subsistence, which must result in misery and undernourishment with all the evils which follow from these conditions. Starvation, disease, famine, and war are the ultimate checks on population which cannot be evaded unless something is done to restrain the fecundity of the people. But we must note that Malthus speaks of another way of escape — through "moral restraint," or control of the birth rate. The checks to population may be classed under two general heads: the preventive and the positive checks. The positive checks operate to eliminate life already in existence, the preventive, to prevent new forms from being born. They are "all resolvable into moral restraint, vice, and misery." It was desirable in Malthus' view to increase "moral restraint" in order to decrease "vice and misery."

Criticism of Malthus' theory. This argument of Malthus is open to attack at several points. His ratios of the increase of population and food are hypothetical and lacking in probative value. It has not been proved that population, even under favorable conditions, will inevitably increase in geometrical ratio and continue so to increase through an indefinite period. Malthus here made the fundamental error of confusing a potential with a real increase; he mistook a biological possibility for a human tendency. Even if we grant that the increase is in geometrical ratio, this gives us no clue as to the absolute rate of increase or the comparison of this rate with that of the increase in food. A geometrical ratio may result in so slow a rate of increase and an arithmetical ratio in so rapid a rate, that the one will not overtake and outstrip the other within a period of time brief enough to be of concern to human beings. In later editions of his book Malthus himself was inclined to place less reliance in the ratios as proof of his theory.

But a greater weakness of his theory is that it leaves out of account the possibility of an advance in the arts of production. The comparison of the increase of people with that of food is based on the unacknowledged assumption that no improvement will occur to multiply the productive power of the natural environment. It is assumed that each succeeding generation of people will win their subsistence from the soil with the aid of no arts unknown to their predecessors. This we know is not the case. Rapid progress in the arts of production during the past century and a half has started each succeeding generation of men on a new and higher economic plane. It is conceivable that this increase in productive power may proceed so rapidly as to postpone indefinitely the pressure of increasing numbers on the supporting power of the soil. Safe prediction on this point must be based, not on the rules of logic and the comparison of hypothetical ratios, but on accurate knowledge of the facts of economic progress and of the actual increase in population.

Modified statement of the Malthusian theory. Thus modified the Malthusian doctrine contains a truth of great importance. There is no disputing the statement that so great is the fecundity of the human race that, if no restriction is placed on the birth rate and no great advance in the arts of production occurs, the number will soon reach the limit of the available food supply and be kept within this limit only by starvation and death. Under no system of social organization is it possible to escape the fact that if population expands more rapidly than the means of subsistence, poverty must ensue. High birth rates, high death rates, pressure for employment, and low real incomes are connected phenomena. This results from the operation of the law of diminishing returns in agriculture and in all forms of industry devoted to the improvement of raw materials taken from nature. "It is vain to say that all mouths which the increase of mankind calls into existence bring with them hands. The new mouths require as much food as the old ones, and the hands do not produce as much."[1] On a given stage of the arts, there is a certain density of population which

[1] J. S. Mill, *Political Economy*, bk. I, ch. xiii, par. 2.

will be the most efficient from the point of view of return per head. If population has not attained this number, every additional man will increase the productive effectiveness of the group and raise the average income. On the other hand if it passes beyond this number, each new producer will add a smaller amount to the social product and the average income will fall. Considering that human increase is an almost automatic biological process with vast powers of expansion, and that progress in the arts of production is continuously retarded by the limitations of human intelligence, the race of men is at all times confronted with the possibility of a redundancy of numbers. Stating Malthus' doctrine in more cautious terms, it is true that population tends to increase up to the limit of the supporting power of the environment, on a given stage of the arts. This tendency must inevitably cause suffering, misery, and vice (the positive checks of Malthus) unless (1) it is held in check by control of births (the preventive check) or (2) its effects are postponed by timely advances in science and the arts of production.

Nineteenth century expansion. During most of the nineteenth century the lesson which Malthus sought to drive home was obscured by the fact that, although population was multiplying freely, life was getting easier. In the course of the century Europeans much more than doubled in number and yet were better fed than at the beginning. The explanation is to be found in the tremendous improvement in the arts known as the Industrial Revolution, to which we have already given some attention. (See especially Chapter IV.)

In England the Industrial Revolution had only begun at the time of Malthus. Its disturbing influence contributed to the misery and discontent which Malthus saw around him. The spinning jenny came into use in the same year in which the Essay was first published. Cartwright's loom began to be used in 1801. But it was not until 1828 that the first commercial steamer crossed the Atlantic, and not until about 1870 that the full effect of invention and international trade had worked itself out in the world-wide division of labor. Goods can now be brought from the most distant countries more cheaply and almost as quickly as they could be carried

from London to Cornwall in the time of Malthus. The population of Great Britain and Ireland was 16,000,000 in 1801 and 41,500,000 in 1901. Total British imports and exports were £37,000,000 in 1791 and £870,000,000 in 1901. The population problem with which Malthus was especially concerned, the problem of feeding a rapidly increasing number of Englishmen on the produce of an island which remained the same in size, was solved, for a hundred years at least, by an immense increase in the production of manufactures and the exchange of these for food and raw materials from new continents. As numbers increased food actually became cheaper. During the latter half of the nineteenth century, a unit of labor applied to industry in Europe could be exchanged for a steadily increasing quantity of food.

This remarkable increase in population, permitted by the great advance in the arts, was not confined to England. During the nineteenth century the population of the world increased by approximately one hundred per cent. The number is usually given as about 850,000,000 in 1830; at the present time the best estimates place it at approximately 1,700,000,000. The figure for the earlier period is of course in large measure an estimate; even at the present time only about one half is known by actual count.

The rate of increase during the century was very unequal in the different grand divisions: in millions, Europe increased from 200 to 465, Asia from 400 to 870, Africa from 100 to 140, and the Americas from 20 to 205.

POPULATION OF CERTAIN EUROPEAN COUNTRIES IN MILLIONS[1]

	1801	*1901*	
European Russia	40.1	112.8	
France	26.8	38.6	
Germany	25.0	56.4	(1900)
Austria-Hungary	25.0	45.4	(1900)
Italy	17.5	32.5	
Great Britain & Ireland	16.3	41.5	
Spain	6.0	18.1	(1897)

Over-population still a menace. It would seem at first sight, in view of these facts, that the forebodings of Malthus could have

[1] E. B. Reuter, *Population Problems*, p. 93.

little relevance to the problems of the modern world. Fundamentally however the issue remains the same. How is the population of the world to be restrained from increasing faster than the world's food supply, except by checks on population? Many modern students of this problem are convinced that the conditions which made possible the unprecedented expansion of the European peoples during the past century are passing, never to return. The menace of over-population has again become a problem of grave practical importance.

It is very doubtful whether the world will experience again a rapid multiplication of its food producing power similar to that which has occurred during the past hundred and fifty years. Some improvement may be made in agricultural implements and in transportation. But the amount of arable land drawn into use has now virtually reached its maximum. Efforts to increase the food supply must be concentrated on a fixed land area, taking the form of greater intensity of cultivation, with its consequence, a lower average output per worker. This condition will become increasingly apparent in the more highly industrialized nations as the increase of population in undeveloped countries absorbs a larger proportion of the foodstuffs which these undeveloped countries produce. Already Europe fails to produce enough food to feed her people. The New World from which Europe draws the remainder of her food supply is increasing rapidly in population, and it is also becoming industrialized so that it is less desirous of accepting manufactures from Europe in exchange for food. It seems inevitable that nations which are not now self-supporting as regards food and raw materials must experience either a decline in the standard of living or a reduction of population to a point more nearly comparable with their natural productivity.

The total population of the earth at the present time is generally given as approximately one billion seven hundred millions. The total land area, exclusive of the polar regions, is something more than fifty million square miles. But the habitable part of the globe is very much less than its total land area, and not above one fifteenth of the habitable regions of the world is under cultivation.

Belgium is perhaps the only country of the modern world cultivating so much as one half of its total area. Canada and Asiatic Russia stand at the opposite extreme. These are regions of immense fertility, but in neither case with so much as one hundredth of the total acreage under cultivation.

AREA AND POPULATION OF THE EARTH BY CONTINENTAL DIVISIONS [1]

Continental Divisions	*Area in Square Miles*	*Population*	*Density per Square Mile*
Asia	17,052,000	890,000,000	52
Europe	3,821,000	475,000,000	125
North America	8,040,000	145,000,000	18
South America	7,018,000	61,000,000	9
Africa	11,605,000	140,000,000	12
Australia and Oceania	3,457,000	9,000,000	3
Polar Regions	4,892,000	.....	...

How fast can population grow? The theoretical power of increase of all forms of organic life is very great. Certain minute organisms, of which the bacteria and certain disease bacilli are typical, reproduce themselves at a rate so rapid that no conception can be formed in terms of number. By a simple process of growth and division the numbers increase in a geometric ratio. Assuming an hourly rate — and many forms exceed this — the descendants of a single individual would exceed ten million in number before the end of the first day. To consider just one other case of the power of fecundity among the lower forms of life, Darwin computed the reproducing power of the elephant. This is the slowest breeder of all animals, yet from a single pair there would be ninety million living descendants after a lapse of 750 years if the normal rate of increase were unchecked.

In spite of the long periods of incubation, childhood, and immaturity and of the relatively short fertile period of the female's life, the human species nevertheless has the physiological capacity to double its numbers in about seventeen-year periods. This is slow as compared with the lower forms, but it is all-sufficient to people the earth with unnumbered millions within a historic epoch. "Assuming a doubling in twenty-five-year periods to be well within the

[1] E. B. Reuter, *Population Problems*, p. 18.

historic as well as the physiological limit, the descendants of a single pair living at the time of Christ would today be sufficiently numerous so that the entire surface of the earth would furnish standing-room for about one eleventh of their number. Assuming that they doubled in numbers in fifty-year periods, the number of descendants of the single pair would today approximate one thousand times the present population of the globe."[1] Even at the present average rate of increase of the population of the world — a rate of increase which is obviously everywhere severely restricted — a single pair would produce in 1,750 years descendants equal in number to the present population of the world.

Of the nations regarding which we have reliable statistics the highest birth rates are to be found in Russia, British India, and French Canada. Here the average birth rate is 50 per thousand while in some localities it mounts as high as 55 or 60. A population increasing with this speed is potentially capable of doubling every 25 years, or expanding in a century to 16 times its present size.

Checks to population increase. The tendency of man and other animal forms to increase in numbers can nowhere find unlimited expression for any considerable period of time. Certain checks are always and everywhere either actively or potentially in operation, and these set the upper limit to the number of the population. As stated above, they may operate either to eliminate life already in existence or to prevent new forms from being born. In other words, since the growth of population depends on the excess of births over deaths, the rate of expansion can be reduced either by diminishing the number of births or by increasing the number of deaths.

Positive checks. The positive checks to the increase of plant and animal life are obvious. These lower forms of life quickly exhaust the supporting power of the soil and kill each other by over-crowding. But the operation of these checks is less obvious in the case of human beings, for several reasons. In the first place man is not limited in his search for food to his immediate environment but may move thousands of miles if necessary in search of more favorable conditions. Moreover man is continually improving

[1] E. B. Reuter, *Population Problems*, p. 115.

his methods of winning subsistence from the resources of nature. Despite these advantages however man's increase in numbers is subject to the same ultimate check. He relies on the earth not only for his food but for every other resource which he uses to satisfy his needs. All natural resources are exhaustible, and increasing pressure of human life upon the earth must eventually press these resources toward the point of exhaustion.

Chief among the positive checks to population growth is lack of the means of subsistence for the individuals born. If growth of numbers is not checked by other causes, it will ultimately be checked by the sheer impossibility of obtaining food necessary to maintain the life of the increasing number of units. The food supply thus operates inexorably in setting the upper limit to the world's population. It does not prevent the birth of new individuals, but it fixes the number of such individuals that can survive. If births exceed this number, starvation results and a balance between population and means of subsistence is again established. In extreme cases this check on the increase of life takes the form of famine. Famines seem to have been very common among primitive peoples, much more so than among more advanced peoples, because of the low stage of the industrial arts and the meagre facilities for communication, transportation, and exchange. Among the ancient civilized peoples there occurred many severe famines, while famines continued to be a common affliction in Europe down to the middle of the seventeenth century and even occurred during the eighteenth. In the nineteenth century there were many serious famines in Russia and in India, China, and other countries of the Far East. It is often said that previous to the nineteenth century there was never a decade in which famines did not occur.

Other important forms of the positive check are floods, pestilence, and war. Even so apparently natural a phenomenon as flood is not unrelated to the question of the density of population. The pressure of population on the environment leads to a sacrifice of the forests which control the flow of surface water. Perhaps the basic cause of the recent disastrous floods in China has been the exhaus-

tion of timber in the countryside bordering on the headwaters of its great rivers. Pestilence also increases with population pressure, because of the greater difficulty of sanitation and isolation of the infected individuals and of the susceptibility of an undernourished people to disease. Disease in its manifold forms operates at one remove from starvation to limit growth of numbers. It is the most important cause of death in civilized countries. War also is caused to a large extent by population pressure. Throughout all history, as well as at the present time, a difference in the rates of population growth on two sides of a political frontier has been a potent cause of war.

Preventive checks. The methods among civilized nations of limiting increase are chiefly confined to those which reduce fertility, with the exception of the practice of abortion, which is still of some importance in the lower social classes. Chief among these are celibacy, late or deferred marriage, and the use of contraceptive measures. Further light will be thrown on this topic when discussing the declining birth rate.

The death rate. Within the past hundred years, but especially since the latter part of the nineteenth century, there has occurred a sensational lowering of the death rate among the more advanced peoples. While differing somewhat as to time and extent in the different countries, it has been general throughout the Western World. The great progress in public health and medical science which occurred in the third of a century before the World War had cut down the infant mortality rate among advanced peoples by an amount varying between a quarter and two fifths.

The following table shows the decline by decades per thousand population for eight European countries.[1]

DEATH RATES BY DECADES, 1870–1910, EIGHT EUROPEAN COUNTRIES

	United Kingdom	*European Russia*	*Germany*	*France*	*Austria*	*Hungary*	*Spain*	*Italy*
1870	22.4	....	27.2	24.4	31.9	35.9	30.9	30.2
1880	20.0	....	25.8	22.4	30.7	32.4	30.6	28.7
1890	18.9	35.8	23.9	22.2	28.5	32.8	31.6	26.5
1900	17.6	31.8	20.8	20.7	24.8	26.9	27.9	22.6
1910	14.9	29.8	17.4	19.1	22.2	24.7	23.9	21.1

[1] E. B. Reuter, *Population Problems*, p. 146.

The lowest death rates so far reached in the countries where records are kept are those of Holland, Great Britain, Australia, the Scandinavian countries, and some of the American states. The highest rates are those of Russia and Roumania in Europe and Chili in America. The following table records the death rate in certain countries since 1920.

ANNUAL CRUDE DEATH RATES SINCE 1920

	Deaths per 1,000 population
New Zealand	8.7
Australia	9.9
United States	11.8
Netherlands	12.8
Wales	12.8
Denmark	12.9
Norway	13.2
England	13.7
Sweden	13.8
Switzerland	14.3
Belgium	15.2
Scotland	15.4
Germany	16.2
Ireland	17.6
France	18.4
Italy	19.2
Japan	21.9
Austria	21.9
Roumania	23.4
Bulgaria	26.4
Chili	31.1

A high death rate frequently goes with a high birth rate, because both are characteristics of poverty, ignorance, and proletarian conditions of life. The same conditions that favor excessive fecundity make inevitable a high rate of mortality.

The possible decline in the death rate is rather narrowly limited. If the average of life were seventy years, the death rate would approximate 14 per thousand. In China, with an average birth and death rate of probably 50 per thousand, the average span of life is about twenty years. A society with a death rate of 35 would have an average span of life of under thirty years. If the death rate be reduced to 20, it would mean an average life of fifty years.

If it be reduced to 10, it would mean that every child born would live to be one hundred, or if some die younger, others would need to live longer.

It is evident that a death rate of approximately 9, as in New Zealand, or of approximately 10, as in the Australian Commonwealth, cannot long be maintained. Such rates, in the absence of a large immigration of persons in the safer years of life, are destined to a present rise. It is equally evident that a death rate of 25 or 30 is excessive and that such rates would fall rapidly with the extension of education and the introduction of modern methods of sanitation. The point to which the death rate may be reduced is not definitely fixed. It is probable that the point is not the same in different racial groups. It is also likely that there will always be differences due to climatic and geographic conditions. Every reduction in the death rate beyond a certain point is made at greater cost and effort.

The reduction of the death rate has been especially marked in the case of infant mortality, which is always higher than the mortality of other age groups. "The first census taken by the Japanese in Formosa indicated that the Chinese there lose one half of their children before they are six months old. Some years ago a third of the Russian babies and a fourth of the Bavarian babies regularly failed to live as long as one year. On the other hand in the best Scandinavian or American communities not more than one infant in twenty fails to survive the first year, and in New Zealand, where the baby-saving campaign has been pushed farther than anywhere else, there are cities which lose but one infant in twenty-six."[1] To a large extent this reduction in the rate of infant mortality is due to a lowering of the birth rate. It has been found that each reduction in the rate at which infants are born is accompanied by a reduction in the rate at which they die.

The rate of infant mortality differs in various social classes, being higher in the lower income-groups and lowest among professional people, as may be seen by the following statistics relating to England and Wales.[2]

[1] E. A. Ross, *Principles of Sociology*, p. 30.
[2] Harold Cox, *The Problem of Population*, p. 117.

England and Wales 1911	*Births per 1,000 Married Males under 55*	*Deaths of Infants per 1,000 Born*
Medical practitioners	103	39
Solicitors	100	41
Clergymen, Church of England	101	48
Dock laborers	231	172
Earthenware makers	181	172
Costers and hawkers	175	196

The birth rate. An event which has caused much more discussion than the declining death rate is the remarkable decline that has occurred in the birth rate. About 60 years ago the population of France began to show a sharp decrease in fecundity. The date at which births began to decline in England is usually set at 1878. In this year wide publicity was given to an attempt on the part of the government to punish a group of people for spreading knowledge of birth control. The most important result was an immediate popularization of this proscribed knowledge, with the result that the birth rate fell by as much as one third in the next thirty-five years. From France and England the decline in birth rates spread in the next decades to all countries of Eastern Europe and to Australasia. Even in the United States, despite the flood of immigration, there was a drop of 35 per cent in the proportion of children under five years of age to women of child bearing age.

The following table indicates the degree to which the birth rate has fallen in various European countries:[1]

	Births per 1,000 Population 1871–1880	*1901–1910*
Denmark	31.4	28.6
Norway	31.0	27.4
Sweden	30.5	25.8
Finland	37.0	31.2
Austria	39.0	34.7
Switzerland	30.7	26.9
Germany	39.1	32.9
Holland	36.2	30.5
Belgium	32.3	26.1
Italy	36.9	32.7
Australia	36.1	26.5
New Zealand	40.5	26.8
France	25.4	20.6
England and Wales	35.5	28.2

[1] Harold Wright, *Population*, p. 112.

A comparison of the birth rates of various countries at the present time (since 1920) is given below:

Country	*Birth Rate per 1,000 Population*
France	18.7
Ireland	20.0
Spain	21.4
Scotland	21.7
United States	22.4
Switzerland	22.7
Wales	22.7
England	23.1
New Zealand	23.3
Belgium	23.7
Sweden	23.8
Australia	24.9
Canada	25.1
Denmark	25.8
Norway	25.9
Netherlands	27.8
Finland	28.5
Germany	29.8
Italy	31.3
Austria	31.4
Japan	34.2
Chili	37.2
Bulgaria	40.3
Argentina	42.1
Roumania	46.2

It should be noted that, in spite of the decline in the birth rate since the seventies, the populations of practically all countries experiencing this change have rapidly increased. This is owing to the larger volume of the population to which the rate, though reduced, is applied and to the fact that the death rate has declined even more rapidly than the birth rate. The situation in France, where there is now a stationary if not an actually decreasing population, is due not solely to the very low birth rate but as much to the fact that the death rate has not been lowered there as much as in many other countries.

Causes of the declining birth rate. The motives for limiting the size of the family are many. In quality they run the whole gamut from the most noble, pure, and altruistic, to the most base and selfish. They may be listed briefly as follows: (1) the economic

burden of children, especially in urban environments, because of the high cost of living, poverty, employment of women, child labor laws, the lengthened period of compulsory school attendance, etc., (2) increasing regard for child life, its nurture and education, and a keener sense of parental responsibility, (3) desire for social position, attributed to a selfish love of luxury, desire for comfort and economic security, vulgar ambition, prudence, etc., (4) changing position of women as regards increased liberty and economic independence, dislike of child-rearing, avoidance of the pain of child-birth, (5) the increasing demand for long preparation for the professions and other occupations requiring considerable training, which leads to late marriages; and so on, most of the factors being resolvable into a rising standard of living.

The chief way in which these motives have found effective expression has been through the spread of knowledge concerning the means of birth control. In 1877 Mr. Charles Bradlaugh and Mrs. Annie Besant were prosecuted in England for publishing a book, called *The Fruits of Philosophy*, in which they advocated various devices for preventing conception. The trial was a huge advertisement for Neo-Malthusianism. In 1877 the Neo-Malthusian League was founded and a vigorous propaganda carried on. Propaganda in other countries followed and also prosecutions, as in Australia, India, America, and elsewhere. In America there have been and still are various state and national edicts designed to prevent the spread of birth control information, but in this country as in many others birth control is a regular practice among certain classes and is the most important single cause of the declining birth rate.

It is not part of our present task to go into the pros and cons of the matter. It may be stated however that the main arguments advanced against the use of contraceptive methods are moral or religious and that practically all religious bodies, as far as their opinions can be discovered, are strongly opposed to this method. On the other hand it cannot be doubted that for mankind as a whole declining birth rates and lessening pressure on population mean progress.

The differential birth rate. It is a familiar fact that at the present day the birth rate is higher among certain sections of the population than among others. There is reason to believe that this was also the case in some of the older civilizations, notably in Rome. We may first look at the facts and afterward at the causes.

Many investigations have shown that there is a higher fertility among the lower social grades. Heron found that there is a high average correlation between the birth rate and conditions which indicate a lower social grade, such as the number of pawnbrokers per 1,000 inhabitants, the amount of child employment, pauperism, overcrowding, and other indications of poverty and lack of culture. The following table from Bertillon is accepted by Newsholme as probably accurate so far as the general situation is concerned, though he does not consider that the specific figures are entirely trustworthy:

ANNUAL BIRTHS PER 1,000 WOMEN OF FERTILE AGE IN DIFFERENT ECONOMIC STRATA

Classification	*Paris*	*Berlin*	*Vienna*	*London*
Very poor quarters	108	157	200	147
Poor quarters	95	129	164	140
Comfortable quarters	72	114	155	107
Very comfortable quarters	65	96	153	107
Rich quarters	53	63	107	87
Very rich quarters	54	47	81	63

Whether or not a differential birth rate results in a differential contribution to the next generation depends on whether there is a differential survival rate and, if so, whether or not it compensates for the differences in the birth rate. All available evidence goes to support the conclusion that the higher the social status the lower is the fertility and that this initial difference is not removed by subsequent differential mortality. This is apparently true not only for England but also for other countries in which the economic situation is similar, and it is a relatively recent phenomenon.

In this country studies in the field of the declining birth rate have been confined pretty much to college men and women. It has been found that the average Harvard graduate is the father of three fourths of a son and the average Vassar graduate the mother of one

half of a daughter and that the average family of American men of science is only 2.22 as compared with an average of 4.66 for the country. Many statistical studies of Yale, Harvard, and other educational institutions disclose similar results.

Another outstanding characteristic of the falling birth rate has been wide variation between different sections of the country. The New England section seems to have been the first affected and the one where the decline has been greatest. But other highly industrialized sections of the North and East show the same tendency in only slightly less degree. The agricultural South and large sections of the West have shown a relatively small decline.

The difference is likewise marked as between urban and rural populations. The birth rate is higher in the country and the families are larger than in the city. The rate has declined in both rural and urban sections, but the decline appeared first in the city populations, and there it has proceeded farther. Economic differences account largely for the differences in the decline, though impediments to the spread of birth control information in the rural areas is also an important factor.

In immigrant-receiving countries like the United States, the differential birth rate has further significance, in that the immigrant stock is more prolific than the native American. The investigations of the Immigration Commission showed the average family of native white women of native parentage, married from ten to twenty years, to be 2.7 children, while 4.4 children was the average for white women of foreign parentage. The difference in fertility was nearly two to one. For the same age groups, women under forty-five years of age and married from ten to twenty years, 7.4 per cent had no children. The percentage of women of native parentage having no children was 13.0; the percentage of foreign-born women having no children was only 5.7. Forty per cent of the native-born married women had only one or two children; nineteen per cent of the foreign-born women had one or two children. Ten per cent of native women of native parentage had five or more children; thirty-three per cent of the foreign-born women had five or more children.

"This situation — a differential birth rate — may produce real race suicide, in the sense in which the term was originally used by its inventor, Professor E. A. Ross. That is to say the native stock may so decline relatively that the foreign element will become greater than it in numbers, owing simply to differences in the rate of increase. . . . There is much evidence that the best blood of America is being thinned out by the exercise of a conscious limitation of births. If the native stock deliberately and increasingly limits its own increase in this way and if at the same time the foreign stock exercises a much lesser check on increase of numbers, in time the obvious result of this differential birth rate will be to change the character of the predominant racial type in America."[1]

The question of quality of population. The differential birth rate has been widely discussed, and there has been some speculation concerning its effects on the quality of the population stock. It is obvious that the groups with the highest birth rates must, unless their increase is checked by a high death rate, gradually crowd out the less fecund groups. Furthermore if the expanding groups are possessed of inferior characters of a sort transmissible by heredity, the result must be a gradual degradation of the quality of the population.

The significance to population quality of a differential birth rate between racial stocks turns upon the extent to which one racial stock can be shown to be innately superior to another. The trend of scholarly opinion in recent years has been to minimize the differences in racial ability popularly assumed to exist. But there are in general great economic and social differences between the later immigrant stocks and those of native parentage. The social and cultural effects of the excessive increase of the foreign-born stocks are therefore unfortunate, whatever may be the result of their increase on the innate quality of the population. Very little scientific attention has been paid by students of population problems to the quality of the population. Qualitative problems have assumed an importance only since the publication of Darwin's

[1] M. R. Davie, "Immigration and the Declining Birthrate," *The Scientific Monthly*, July, 1924, Vol. XIX, p. 76.

Origin of Species in 1859. Further impetus was given to this side of the population problem by the work of Sir Francis Galton, who coined the word "eugenics." Today this word is associated with a movement to increase the quality of the population by increasing the birth rate of the superior family strains and reducing that of the inferior. The movement, in other words, has both a positive and a negative side.

Eugenics. On its positive side, eugenics rests on certain assumptions which are not entirely tenable, and the student should not overlook certain fundamental truths, which may be briefly stated here.

The classes with lowest birth rates are not necessarily the biologically best. There is need to discount the class bias that so often leads to the tendency to identify the socio-economic, with the psycho-physical *élite*.

The eugenic doctrine involves the assumption that exceptional capacity is a family trait rather than an individual variation. The extent to which this is true is a matter that has not as yet received sufficient scientific investigation.

There is frequent failure to distinguish the cultural from the biological concepts involved. Race improvement and race deterioration are physical, biological concepts; civilization — the progress or decline of culture — is a social concept. Social and cultural and national progress are not matters of biology and not necessarily contingent upon race improvement; a biological improvement in the race is not a prerequisite to social and cultural advance. It may be said further that while talent will not be created by opportunity, opportunity is essential if talent is to develop and function. The person is the product of inheritance and environment. One sets the possibilities of development, the other determines the direction and the extent to which the possibilities are to be realized. Both must be considered.

Not only is there the problem of discovering just who are the "fit"; there is also the question as to what we should breed for. Even if we knew how to produce children with certain characteristics, we are not yet agreed as to what characteristics we want. As

Dean Inge says: "The two ideals, that of the perfect man and that of the perfectly organized state, would lead to very different principles of selection. Do we want a nation of beautiful and moderately efficient Greek gods, or do we want human mastiffs for policemen, human greyhounds for postmen, and so on?" Many of us would answer that we do not want either the one or the other; we would much rather have the present varieties of human beings.

Negative eugenics, on the other hand, presents a more definite and practical program. Its object is to stop the increase of the biologically unfit, especially where the defect is due to heredity. Measures such as segregation, sterilization, prohibition of marriage, and the like have been suggested and in certain respects put into practice. Most communities now have laws denying the right of marriage to certain classes of defectives, such as the feeble-minded, insane, and epileptic; most have hospitals and institutions where such persons may be confined; while some have laws providing for the sterilization of such individuals. These laws are not adequately enforced, nor will they be until states come to make an examination of the total population and a registration of all defective members. We do not yet know all those who are the "unfit." Nevertheless there is a large practical problem presented here, and the question is bound to arise in connection with the regulation of immigration, the problem of national vitality and military strength, the question of the differential birth rate, and the subject of race suicide. Certainly the best way of getting rid of an inferior population is to prevent the inferior people from increasing.

CHAPTER XLIII

IMMIGRATION

Spread of population. Immigration is simply the most modern phase of the immemorial tendency of men to migrate in order to improve their conditions of life. Mankind has ever been on the move. In prehistoric times these movements of population took the form of more or less blind dispersions or wanderings, in which primitive tribes scattered about in every direction in search of food, hunting grounds, etc. The most important types of migration which history has recorded are invasion, conquest, colonization, and immigration. The last two are the most recent, dating back only to the Discoveries Period which opened up the outlying continents. The Europeans were the first to take advantage of the new and seemingly unlimited areas thus discovered, and they have since become spread over most of the world. While the forms of migration in all these instances differ, they have this in common — the desire for more land, which the late Professor Sumner termed "earth hunger." Since all means of subsistence are derived ultimately from the land, more land has meant better conditions of life. This is true whether the chief method of self-maintenance is by hunting, cattle-raising, agriculture, or industry. It has been the desire for more land which has caused the more energetic races of mankind to spread over all the inhabitable portions of the globe.

The movement in every case has been away from countries that are overpopulated and toward countries that are underpopulated. The most fundamental controlling condition of the status of any group of people is the ratio of their numbers to the land at their disposal. If, on a given stage of the arts, the population is low in proportion to the amount of land, life is easy and the competition of man with man is weak. When more persons are

trying to live on a square mile than it can support on the existing stage of the arts, life is hard and the competition of man with man is intense. In the former case the average condition will be one of comfort and the population will all be near the average, while in the latter case the average condition will not be one of comfort and the population will cover wide extremes of comfort and misery. In an overpopulated country the extremes of wealth and luxury are presented side by side with the extremes of poverty and distress. They are equally the products of an intense social pressure.

In overpopulated countries, like those of Europe and Asia, not only is the competition of life intense and the general standard of living low, but as a result of the same social pressure military burdens are heavier, liberty is restricted, and there is likely to be greater political oppression and less religious tolerance. In new and relatively underpopulated countries on the other hand conditions are much more favorable. If there is abundance of land and few men to share it, the men will all be equal. Each landholder will be his own tenant and his own laborer. Social classes are lacking. Wages are high and the mass of men is well off. Herein we find the motives for immigration. The old and overpopulated country is repellent with respect to economic, political, religious, and social conditions; the new and underpopulated country is attractive in these ways. Hence the movement of population has been from the former to the latter. The country which receives the stream of immigration is newer, much less thickly settled than the other, and living conditions are more favorable. It is more democratic, its people enjoy greater social and political equality, there is more individual freedom of conduct, and there are fewer traditional and legal restraints, military burdens are lighter, and there is greater latitude for religious belief and practice.

Europe, with a density of population of about 125 persons to the square mile, and Asia, with a density of about 55, have been the main sources of immigration. As matters now stand, Europe and Asia, with two fifths of the world's area, support four fifths of

the world's population; hence the strong tendency to emigrate. The immigrant-receiving countries on the other hand are relatively underpopulated. The greatest of these is the United States, with a density of population of only 35.5 to the square mile; to this country have come since 1776 over thirty-five million immigrants. Other countries opened up by the Discoveries Period to which immigrants have been going are Canada (density of population 2.2), parts of South America (density 7), Australasia (density 2.3), and South Africa (density of Africa as a whole 11).

These are the facts which, since the great geographical explorations of the fifteenth and sixteenth centuries, have spread the population of the European nations over the globe. The most enterprising nations seized the advantage first and have pushed it farthest. The movements of population have been accelerated by all the inventions which have facilitated transportation and communication.

Immigration, which is the most recent form of movement of population, may be defined and its distinguishing characteristics pointed out briefly as follows: Immigration is a movement of people, individually or in families, acting on their own personal initiative and responsibility, without official support or compulsion, passing from one well-developed country, usually old and thickly settled, to another country well developed in the sense of culture, but usually new and sparsely populated, with the intention of residing there permanently.

Immigration to the United States. The history of immigration to the United States may be divided into two major periods: (1) the period of the so-called old immigration — from the beginning of our independent national existence to 1882 — during which time our immigrants came chiefly from the British Isles, the Scandinavian countries, and Germany, and (2) the period of the new immigration, from 1882 to the present, in which most of the immigrants have come from Italy, former Austria-Hungary, Russia, and other countries in the south and east of Europe. This shifting of the main sources of immigration from northern and western Europe to southern and eastern Europe is the most out-

standing fact in the history of the movement to this country. It has meant, in general terms, the coming of peoples who are more alien to the native stock as to race, language, history, religion, customs, and traditions.

A line drawn across the continent of Europe from northeast to southwest, separating the Scandinavian Peninsula, the British Isles, Germany, and France from Russia, former Austria-Hungary, Italy, and Turkey, separates countries not only of distinct races but also of distinct civilizations. It separates countries of representative institutions and popular government from those accustomed to absolute monarchies; it separates lands where education is universal from lands where illiteracy predominates; it separates manufacturing countries, progressive agriculture, and skilled labor from primitive hand industries, backward agriculture, and unskilled labor; it separates an educated, thrifty peasantry from a peasantry scarcely a single generation removed from serfdom; it separates Teutonic races from Latin, Slav, Semitic, and Mongolian races. When the sources of American immigration are shifted from the Western countries so nearly allied to our own to Eastern countries so remote in the main attributes of Western civilization, the change is one that should challenge the attention of every citizen.

Causes of the changed character of immigration. It is impossible within the limits of this chapter to go deeply into the reasons for this changed character of immigration, but some of the main causes are as follows: The great industrial expansion of the United States has taken place since the eighties, and a study of the facts of the case will show that it has gone hand in hand with the increase of immigration from southern and eastern Europe. We used to be chiefly an agricultural nation; we are now a great manufacturing and commercial nation. While we had free land and while the element of skill was a large factor in our industries, conditions were favorable for the immigration of northern and western Europeans. Now that the free land has practically disappeared and the invention of machinery has reduced the need for skilled men and greatly increased the demand for unskilled

workers, we need and have been getting a new type of immigrant. Various other factors, such as the removal of religious and political causes of emigration from northern and western Europe and the fear of competition with recent immigrants of lower standards of living, help to account for the changed character of immigration, but from an economic standpoint the fundamental reasons seem to be the ones just mentioned. Conditions among the skilled workers of northern and western Europe are so little different from conditions of such workers here that there is no motive from that source for emigrating. Our large demand is for unskilled workers, and we are getting mostly unskilled immigrants, no matter from what section of Europe they come. Southern and eastern Europe can furnish proportionately more unskilled industrial workers than northern and western Europe, because the general standard of living is lower and there is less industrial development; so most of our immigrants come from that section.

Restriction. The changed character of immigration has had more influence than any other single factor in developing a body of opinion in favor of restriction of immigration. This has recently found expression in the percentage restriction laws, which constitute a radical departure from our former policy of regulating European immigration. Throughout the period during which we have legislated on the subject of immigration, three different types of laws have been enacted. These involve the principles of exclusion, selection, and restriction.

Exclusion. The principle of exclusion has been applied solely to Asiatics. The Chinese were the first to be excluded, and they have been kept out by a series of acts beginning with that of 1882, which suspended the immigration of Chinese laborers. At Japan's request we entered the so-called "Gentlemen's Agreement" in 1907, a passport arrangement which provides that Japan will not, except under certain conditions, issue passports good for continental United States to laborers. Other races have been excluded by the immigration act of 1917, which created the Asiatic barred zone from which no immigrants may come. This area, defined by degrees of longitude and latitude, includes India, Siam,

Indo-China, parts of Siberia, Afghanistan, and Arabia, and the islands of Java, Sumatra, Ceylon, Borneo, New Guinea, Celebes, and various smaller groups. This is the full list of countries with which we have so far decided to deal on the principle of exclusion. It should be noted however that exclusion applies in all these instances only to the laboring classes. Travelers, students, teachers, and officials, as well as certain other exempt classes, are still free to come.

It will be noted that we have applied the principles of exclusion solely to the darker-skinned races. This has been done largely on racial grounds, although the very low standard of living of these people, which makes it impossible for American workmen to compete with them, has also been a prominent factor. Furthermore these races are so diverse that assimilation is impossible. We believe that our best interests will be served by keeping the United States as far as possible a white man's country. We cannot assimilate the yellow, brown, and black races, and we should only be inviting trouble and adding to our already large and serious race problems by admitting members of such races.

Selection. The principle of selection has been applied to European immigrants since 1882, when the first inclusive immigration law was passed by the Federal Government. Congress was induced to legislate on the subject because with free and unregulated immigration certain evils had grown up which demanded attention. It was obvious to everyone that we could not go on admitting all manner of persons without getting most of the dregs of Europe. Indeed many European nations had been making a practice of using the United States as a dumping ground for their criminals, paupers, and other undesirables. In self-defense then we had to regulate immigration, and the principle of selection, which imposes qualitative tests for admission, was adopted. The law of 1882 prohibited the entry of convicts (other than political offenders), idiots, lunatics, and persons likely to become public charges. Since then we have had a series of laws of increasing severity, which have added to the list of ineligibles and have become more and more stringent regarding the deportation of those

who should have been excluded but who slipped by. The most recent general immigration act, that of 1917, applies a test of physical, mental, moral, and, in a limited way, educational fitness and to this extent is a selective law. Its purpose is to exclude from admission all aliens who upon examination at our ports are found to fall below the prescribed physical, mental, or moral standard, those who are diseased, and (with some exceptions) those who cannot read in the English language or some other language or dialect.

Restriction. Without doubt the feeling has been growing in recent years and has received a great impetus from our World War experience that we need to go beyond the selective principle and find some way to restrict immigration; that is, to reduce greatly the number of aliens that are to be admitted in any year. Restriction, be it noted, presents the quantitative view and would keep out many more than have been refused admission under the selective laws. There are many arguments in its favor. An important one is the recent volume of immigration, which, except for the interruption caused by the World War, has been running an average of about a million a year. Of course some immigrants return, especially during hard times, but after allowing for that there is still a well-founded feeling that too many aliens are coming and that they are coming too fast to be assimilated. It is the proportion of native stock to foreign stock that determines, roughly, the ability of the former to assimilate the latter. Yet statisticians report that two thirds of those who now comprise the American nation are of foreign birth or with foreign parentage in one or both parents. That leaves but one third of the nation to lay claim to native birth with full American parentage. So it appears that assimilation to the American type is becoming more and more remote. Recent immigration, with its great diversity of foreign elements and its huge volume, makes the problem of assimilation more difficult. The tendency of immigrants to concentrate in the cities, forming foreign colonies, adds to the problem.

Experience in the World War brought home the fact that we had not proceeded very far in the matter of assimilation. It made us

notice, seemingly for the first time, the facts that about five of the fourteen million persons of foreign birth in this country cannot read, write, or speak the English language; that over half of the foreign-born white males of voting age were not citizens; that European governments were in some instances inducing their subjects here to retain their old allegiance; and that in some cases those immigrants whom we had thought to be the most completely assimilated were not assimilated at all. The "melting pot" was discovered to exist chiefly in imagination. This evidence of non-assimilation, and much more that is not mentioned, is now receiving considerable attention. It seems but fair to assume that some restriction of immigration is necessary if we are to effect any real assimilation.

Another argument in favor of restriction is that of race suicide, in the original meaning of the term. The lowest birth rate is to be found among the native stock, the highest among the recent immigrant groups. If these unequal rates of increase continue, this nation will lose its former character and from being primarily Nordic in stock will become Alpine and Mediterranean. Other arguments stress the competition of recent immigrants of low standards of living with native workmen, the increased problems of crime and pauperism, and the difficulties of government.

In addition there is the fact that if immigration is not restricted, the general level of conditions here will be reduced to that which prevails in Europe. The birth rate and immigration together have greatly increased our population and removed some of the advantages we formerly had when underpopulated to a greater degree. We are still underpopulated as compared with Europe, but it must be remembered that overpopulation and underpopulation are only relative terms and that as time goes on any underpopulated nation, owing to the constant and universal tendency of population to outstrip the means of subsistence, is surely moving forward toward the other status and is speedily losing its natural advantages. There is evidence that that is taking place in the United States. We still have many advantages over European nations however and it is most certainly to our interest to maintain

those advantages. Unrestricted immigration, with a huge influx of persons of low standard of living, will increase the pressure in this country and finally reduce us to the lowest level of European countries that are permitted to send us immigrants. If immigration were allowed to work out its final result, standards here would be reduced to the level that prevails in Europe, and immigration would then automatically cease, for conditions here would be so much like those in Europe that there would be no incentive to emigrate. We need restriction of immigration to guard against such an event.

Restriction acts. The growing sentiment in favor of restriction finally found expression in the temporary act of 1921, which presented a totally different type of regulation. It restricted immigration on a basis wholly numerical, it was devoid of selective features, and its administration was largely an arithmetical problem. This act, which went into effect June 3, 1921, provided that the number of aliens of any nationality admissible to the United States in any fiscal year should be limited to three per cent of the number of persons of such nationality who were resident in the United States according to the census of 1910, it being also provided that not more than twenty per cent of any annual quota could be admitted in any one month.

This act expired by limitation at the close of the fiscal year, June 30, 1924. It has been supplanted by the permanent quota limit act, passed May 26 of that year and officially known as the "Immigration act of 1924." Under the act of 1924, "the annual quota of any nationality shall be 2 per centum of the number of foreign-born individuals of such nationality resident in continental United States as determined by the United States census of 1890, but the minimum quota of any nationality shall be 100." Not more than ten per cent of any annual quota may be admitted in any month except in cases where such quota is less than 300 for the entire year. For the purposes of the act, nationality is determined by country of birth.

The quota area, under the act of 1921, was limited to Europe, the Near East, Africa, and Australasia. The countries of North

and South America, with adjacent islands, and countries from which immigration was otherwise regulated, such as China, Japan, and countries within the Asiatic barred zone, were not within the scope of the quota law. Under the new act however immigration from the entire world, with the exception of the Dominion of Canada, Newfoundland, the Republic of Mexico, the Republic of Cuba, the Republic of Haiti, the Dominican Republic, the Canal Zone, and the independent countries of Central and South America, is subject to quota limitations.

The effect of using the census of 1890 as a basis, instead of the census of 1910 as in the former act, and of reducing the rate from three to two per cent is obviously to restrict the number of admissible immigrants. The act of 1921 allowed a total annual quota of 357,803; the act of 1924 allows only 164,667. After July 1, 1927, the total annual quota will be further reduced to 150,000, which will be the permanent provision, and the individual quotas will be adjusted on that basis. A different basis of apportionment of the quotas of the separate countries will be used. The total quota of 150,000 will be apportioned among the countries to which the act applies according to their relative contribution to the American population as enumerated in 1920, with the proviso that the minimum quota of any nationality shall be 100. To determine such contribution is a very intricate problem, involving a study of the American people from the earliest days to the time of the last census. The process of apportionment may be outlined as follows: First ascertain the total number of persons in continental United States in 1920 whose origin by birth or ancestry is attributable to the countries included under the quota act. "Such determination shall not be made by tracing the ancestors or descendants of particular individuals, but shall be based upon the statistics of immigration and emigration, together with rates of increase of population as shown by successive decennial United States censuses, and such other data as may be found to be reliable." Secondly determine what fraction of this total number has been contributed by each of these countries. Then apply this ratio to the total annual quota of 150,000.

This will give the quota for each country included. If, for example, 13 per cent of the people in the United States in 1920 having such national origin as above explained were German by birth or extraction, then 13 per cent of 150,000, or roughly 20,000, will be the number of immigrants annually admissible from Germany.

In our immigration legislation the term "immigrant" includes in general those aliens who come here to take up permanent residence and to find employment. According to the law of 1924, immigrants are divided into two classes, quota immigrants and non-quota immigrants, meaning in the first instance aliens who are chargeable against the quotas of their respective countries and in the second immigrants who may enter the United States without reference to quota limitations. The non-quota immigrants include:

(1) An immigrant who is the unmarried child under eighteen years of age, or the wife, of a resident citizen of the United States who files a petition stating that he is able to and will support the immigrant if necessary to prevent such immigrant from becoming a public charge;

(2) An immigrant previously lawfully admitted to the United States, who is returning from a temporary visit abroad;

(3) An immigrant who was born in the Dominion of Canada or in any one of the other countries of the New World mentioned above, and his wife, and his unmarried children under eighteen years of age, if accompanying or following to join him;

(4) An immigrant who seeks to enter the United States solely for the purpose of carrying on the vocation of minister of any religious denomination, or professor of an institution of learning, and his wife, and his unmarried children under eighteen years of age, if accompanying or following to join him;

(5) An immigrant who is a bona fide student at least fifteen years of age and who seeks to enter the United States solely for the purpose of study at an accredited educational institution.

All immigrants must be in possession of an immigration visa issued by a United States consul before they can be admitted to the United States, while the annual and monthly limitation under

the various quotas is controlled through limiting the number of quota immigration visas issued in any month or year. In other words, the quotas are counted or controlled in American consulates, usually in the country where the applicant resides, rather than on arrival at a United States port, as was the case under the former quota limit act, thus obviating the unhappy experience of the three years that law was in effect, when thousands of aliens were brought to the United States in excess of quotas, only to be returned to the country of origin.

Preference in the issuance of quota visas is given (1) to a quota immigrant who is the unmarried child under twenty-one years of age, the father, the mother, the husband, or the wife, of a citizen of the United States who is twenty-one years of age or over, and (2) to a quota immigrant who is skilled in agriculture, and his wife, and his dependent children under the age of sixteen years, if accompanying or following to join him. A preference in the case of persons skilled in agriculture is not applicable to immigrants of any nationality the annual quota for which is less than 300, and in no case shall the combined preferences exceed fifty per cent of the annual quota of any nationality.

An important provision of the act of 1924 is to be found in that section which provides that " no alien ineligible to citizenship shall be admitted to the United States." The import of this provision will be readily understood when it is considered that our naturalization laws state that the provisions thereof " shall apply to aliens being free white persons and to aliens of African nationality and to persons of African descent." This in effect means that persons other than members of the white or the African black race are not eligible to citizenship through naturalization and therefore, with certain exceptions, are not eligible for admission to the United States as immigrants. Included in the category of persons ineligible to citizenship are the Chinese, Japanese, East Indians, and other peoples indigenous to Asiatic countries and adjacent islands. This provision tends to make more effective the exclusion of Asiatics, which had already been provided for by other means, and it will probably minimize the administrative difficulties which have

always existed, especially under the Chinese exclusion acts.[1] The provision, however, greatly offended the sensibilities of the Japanese people.

The provisions of the Immigration Act of 1924 are in addition to and not in substitution for the provisions of the immigration laws. In particular, the law of 1917 is still our basic general immigration law, upon which the act of 1924 has been superimposed.

Considerations in favor of the Act of 1924. A great deal has been said and written both in opposition to and in favor of the act of 1924, which makes several very important changes, not only in our immigration policy but also in the administrative machinery of the Immigration Service. The following considerations are urged in its favor:

1. That it checks the immigration of the less desirable southern and eastern Europeans a great deal more than it does the immigration from the north and west of Europe. The number of immigrants allotted to southern and eastern Europe amounts to only 15.5 per cent of the total quota on the basis of the census of 1890 and 16.3 per cent on the basis of national origin (after July 1, 1927), while under the same methods of computation northern and western Europe will be allowed respectively 84.5 per cent and 83.7 per cent of the total quota. In actual figures the situation will be as follows: Until July 1, 1927, the number of immigrants permitted to enter under the quota from southern and eastern Europe will be roughly 25,500 annually and from northern and western Europe about 139,000. After 1927 the permanent annual quotas will be, for southern and eastern Europeans 24,450 and for northern and western Europeans 125,550, making the total of 150,000 as provided in the law. To appreciate fully the differential effect of this measure, it should be noted that from 1910 to 1914 — the most recent normal period of immigration — the average annual immigration from northern and western Europe was about 183,000, while the annual immigration during the same period from southern and eastern Europe was 738,000. The

[1] *Annual Report of the Commissioner General of Immigration*, 1924, pp. 29–30.

restrictive effect of the law applies much more stringently to southern and eastern Europeans.

2. It accomplishes this restriction without making any overt race distinctions. Discrimination is made possible by the fact that, according to the 1890 census, there were many more people in this country who came from northern and western Europe than from southern and eastern Europe, while the same fact applies, with only a little less force, to the computations as to national origin based on the 1920 census, which will be effective after July 1, 1927.

3. The law will materially reduce the number of immigrants admitted, even more so than was accomplished by the three per cent restriction law of 1921. The small annual quotas of 164,667 until the year 1927 and of 150,000 thereafter are to be compared with the huge volume of immigration prior to the World War. It will accomplish what it was designed to accomplish, and the arguments in favor of restriction may be said to be arguments in support of the present law.

4. The administration of our immigration laws will be greatly improved through the provision of immigration visas. This will make possible a certain amount of selection of immigrants abroad and will furnish a means of regulating the flow of immigration without the unfortunate necessity of excluding so many would-be immigrants here. There will be further improvement in the Immigration Service because of the provision of the act which places the burden of proof upon the alien to establish that he is not subject to exclusion under any provision of the immigration laws.

Arguments in opposition to the act of 1924. The chief opponents of the measure when it was being considered were some employers of immigrant labor, steamship companies, and immigrants already here who wanted their friends and relatives to join them. Opponents of this type of regulation, whether they be selectionists or those in favor of restriction but preferring some other method, make the following objections:

1. The law sets up an arbitrary and fixed method of regulating immigration and is therefore not adaptable either to the needs of

this country for immigrant labor or to the conditions in other countries which give rise to the desire to emigrate.

2. The law will present, as did the temporary act of 1921, a much greater administrative problem arising from smuggling and illegal entry of immigrants. Because the law is so drastic, the problem of enforcing it will be more difficult. The per centum limit law of 1921 proved to be a prolific promoter of smuggling and border running in general, especially across the Canadian and Mexican borders, for this method of getting into the United States was resorted to by great numbers of determined Europeans who failed to find, or were delayed in finding, places under the quotas. Closely allied to border running is the illegal entry of alleged seamen who come to United States ports as members of ships' crews and, taking advantage of shore privileges granted under the law, desert the vessels bringing them and remain in the country. The deserting seaman method of gaining illegal entry has been practiced to some extent ever since immigration laws began to interfere with the unrestricted coming of aliens, but under the quota law such violations have grown to alarming proportions.[1]

3. The law discriminates against southern and eastern Europeans on a group basis, and the grounds upon which such discrimination is made are not tenable. The theory behind the law is: (*a*) That such immigrants are undesirable because they come here, not for religious and political reasons like the old immigrants, but almost solely for economic gain. As a matter of fact, economic motives have predominated in all immigration, both old and new, while religious and political motives are not absent in the new immigration. (*b*) That they are undesirable because unskilled. Yet our demand is chiefly for unskilled immigrants. The southern and eastern Europeans are the main labor supply in many industries. Employers are complaining of a scarcity of unskilled labor. (*c*) That they are undesirable because they settle in colonies. But segregation is not peculiar to them; all racial groups are segregated, the Americans as well as the others. (*d*) That they are inferior racially. The recent immigrants are mostly Alpines and Mediter-

[1] *Annual Report of the Commissioner General of Immigration*, 1924, pp.13–23.

raneans, as contrasted with the immigrants from northern and western Europe, who are chiefly of the Nordic stock. Yet it must be remembered that all Nordics are not superior and that all Alpines and Mediterraneans are not inferior. Our immigration legislation should therefore be on an individual rather than on a group basis. The restrictive law is merely quantitative; it does not allow for the selection of the better types from each section.

4. The law does not bring any improvement in type of immigrant; it is purely mathematical and is devoid of selective features. On the other hand there is great need of raising the selective standards. This is so much more important than merely limiting the amount of bad stock that may be coming in, that the quota law should not be regarded as the end of immigration legislation. A better way of securing restriction is through selection that is very stringent. If selective standards should be raised high enough, it would be possible to effect a great reduction in number of immigrants admitted and at the same time to assure us of a better type. This method would enable us to adjust both as to quality and quantity. Judging from the type of immigrant that has been coming in recent years, we need higher selective standards; that is the most important phase of the subject. These should be on an individual basis. If the southern and eastern European immigrants are inferior in physique, intelligence, etc., to northern and western Europeans, then more of them may be kept out by such selective tests. Here is a justifiable way of effecting that end, not merely a device for arbitrary limitation of numbers, which the quota law provides without giving us any better selective standards than we now possess.

Social consequences of the restrictive policy. To some extent the defenders and opponents of the new policy of rigid restriction represent selfish groups whose interests are affected beneficially or adversely by the law. Hence it is not easy to appraise the merits of the policy on the basis of the arguments presented for or against it. But if we disregard the effect of restriction upon special sections of the people and consider its relation to the general welfare, it will appear that the new attitude toward immigration is

basically sound. With the disappearance of the frontier and the appropriation under private ownership of our natural resources, mere increase in numbers offers slight economic advantage to the country. It is much more important to improve the quality of the population than to increase its size. The tendency of all peoples to expand their numbers by natural increase under favorable conditions can be relied upon to cause population growth commensurate with economic opportunity. Increase through the birth rate produces a more homogeneous people and obviates many of the social and political problems which are created by the process of expansion through immigration of aliens.

Many important consequences will flow from the policy of restriction, if this policy is applied for any considerable period of time. Of chief importance perhaps will be its effect upon different labor groups in the country and upon relative wage rates. Native born Americans rarely become unskilled laborers. A drastic reduction of the relative supply of workmen in this group must result from the new immigration policy. This will raise the lowest rates of wages in the country and tend to reduce existing differences in wage rates. Restriction of immigration will also promote transference among the different groups of workers by allowing our system of free education through trade schools, night schools, and the like to reach almost the entire laboring population. Formerly the workmen of alien birth — an important fraction of the whole — were but little influenced by the educational system. This too will promote mobility of labor and work toward greater uniformity of earnings.

The restrictive policy furthermore will have its effect on the organized labor movement. One great obstacle to the spread of this movement heretofore has been the heterogeneity of the population. The native born skilled craftsmen have shown little tolerance for or sympathy with the immigrant workers; they have made no sincere effort to bring the latter within the existing unions or to aid them in forming unions of their own. As a result, trade unions have flourished almost exclusively among the skilled and have shown the narrowness of outlook and selfishness of purpose

to be expected from a movement thus limited in its membership. These conditions will change as the laboring population grows more homogeneous in language, customs, and economic prosperity with the cessation of immigration. The probable effect will be to strengthen the labor movement and profoundly to affect its policy in the direction of a broader and more constructive outlook. The changed conditions will be most noticeable in those branches of industry — *e.g.*, textiles, the mines, the steel industry — which employ unskilled workers and which in the past have been able to rely upon an abundant supply of this type of labor. The increased bargaining power which results from relative scarcity of supply will be reflected not only in the wage rate but in many other phases of the relationship of the workers with their employers.

Suggestions for Further Reading on Part V. The Distribution of Wealth and Income

I. General

Carver, T. N. *The Distribution of Wealth.* New York, 1916
Clark, J. B. *The Distribution of Wealth.* New York, 1900
Hobson, J. A. *The Economics of Distribution.* New York, 1900
Kleene, G. A. *Profit and Wages: A Study in the Distribution of Income.* New York, 1916

II. Rent

Walker, F. A. *Land and its Rent.* Boston, 1883
Hurd, R. M. *Principles of City Land Values.* New York, 1903
Ely, R. T. *Outlines of Land Economics.* New York, 1924
George, H. *Progress and Poverty.* New York, 1880

III. Interest

Fisher, I. *The Rate of Interest.* New York, 1907
Gonner, E. C. K. *Interest and Saving.* New York, 1906
Cassell, G. *The Nature and Necessity of Interest.* London, 1900

IV. Wages

Taussig, F. W. *Wages and Capital.* New York, 1896
Davidson, J. *The Bargain Theory of Wages.* London, 1898
Hamilton, W. A. and May, S. *The Control of Wages.* New York, 1923

V. Profits

Hawley, F. B. *Enterprise and the Productive Process.* New York, 1907
Knight, F. H. *Risk, Uncertainty and Profit.* New York, 1921
Foster, W. F. and Catchings, W. *Profits.* Newton, Mass., 1925

VI. Personal Distribution

King, W. I. *The Wealth and Income of the People of the United States.* New York, 1915
National Bureau of Economic Research. *Income in the United States.* New York, 1921
Bowley, A. L. *The Division of the Product of Industry.* Oxford, 1919
———*Change in the Distribution of the National Income in the United Kingdom.* Oxford, 1920
Dalton, H. *Some Aspects of Inequality of Incomes.* London, 1920
Hollander, J. H. *The Abolition of Poverty.* London, 1896
Hobson, J. A. *Problems of Poverty.* London, 1896
Bureau of Applied Economics. *Standards of Living.* Washington, 1919
Rowntree, B. S. *Poverty: A Study in Town Life.* New York, 1902

VII. Population

Malthus. *An Essay on the Principle of Population,* Sixth edition. London, 1826
Wright, H. *Population.* London, 1923
Cox, H. *The Problem of Population.* New York, 1923
Thompson, W. S. *Population: A Study in Malthusianism.* New York, 1915
Carr-Saunders, A. M. *The Population Problem.* London, 1925
Reuter, E. B. *Population Problems.* Philadelphia, 1923

VIII. Immigration

Fairchild, H. P. *Immigration.* New York, 1913
Jenks, J. W. and Lauck, W. G. *The Immigration Problem.* Fourth edition. New York, 1917
Roberts, P. *The New Immigration.* New York, 1912
Commons, J. A. *Races and Immigrants in the United States.* New York, 1907
Hourwich, I. A. *Immigration and Labor.* New York, 1922

PART VI

THE ECONOMICS OF GOVERNMENT

CHAPTER XLIV

PUBLIC FINANCE

Government expenditures. In the year 1923 there was expended by government in the United States something over ten billions of dollars. The Federal Government's share of this total was about three and a half billions; the forty-eight states spent not quite a billion and a half; and the various local governments in all the states accounted for a little over five billions.[1] This is a huge sum, which is likely to stagger the reader's imagination. Somewhat easier to grasp, perhaps, is the fact that government expenditure in that year took fifteen per cent (more than one seventh) of the total income of the people of the United States, estimated at sixty-seven billions. The year 1923 is selected as being the latest year for which the figures of both government expenditures and national income are available. Other years, immediately before and after, would show similar conditions, though there exists a general tendency for the cost of government to increase from year to year.

Government expenditure is demanded in order that government may perform the functions required of it. There was a time, in the early stages of man's economic and political development, when government functions could be performed without the expenditure of money. When war was to be waged, the men of the tribe were summoned by the chief and required to bring their weapons (readily diverted from hunting to fighting) and other equipment (practically identical for peace and war) and give their services

[1] The exact figures as calculated by the National Industrial Conference Board (*Tax Burdens and Public Expenditures*, 1925, page 34) are as follows:

Expenditures of the Federal Government	$ 3,459,000,000
Expenditures of the state governments	1,450,000,000
Expenditures of the local governments	5,136,000,000
Expenditures of all grades of government	10,045,000,000

in the common cause. The building of a common assembly hut or a stockade about the common fields or a road or bridge could be accomplished in the same way. Even as late as the nineteenth century in the United States, the construction and repair of public roads and bridges were accomplished by many a rural town in exactly this primitive way, by summoning the citizens to come with their horses and wagons, ploughs, scrapers, and shovels, and "work the road." Similarly food, firewood, and other supplies needed by the government were in earlier days obtained by direct requisition upon the people. By such means government functions were performed without the expenditure of money.

The reader will recognize however that this is a description either of primitive conditions or of sporadic cases in modern times. Present day governments do not ordinarily obtain free services from their citizens or direct contributions of materials. Whatever the government requires, either of goods or of services, in order to carry on its functions must be bought and paid for. Public expenditure is thus coterminous with government service and is indeed the regular measure of the amount of government service.

Analysis of public expenditures. In a previous chapter we have taken a general survey of the important functions of government.[1] Public accounting systems are such that it is a very difficult matter to apportion the expenditures of government among these several functions. Certain general truths however stand out with sufficient clearness.

The cost of war. Among all national governments the principal item of expenditure is for defense. The cost of actual warfare is enormous, and while war is in progress its cost absorbs the bulk of the expenditures of government. The Civil War cost the United States Government about three billion dollars during the four years of war, at which time the annual cost of operating the Federal Government on a peace basis was about sixty million dollars. The South African War cost Great Britain over one billion dollars, and the cost of the Russo-Japanese War to Japan was nearly a billion

[1] Chapter X.

dollars. The stupendous cost of the World War makes everything before it look small in comparison. The United States expended in prosecuting this war over thirty billion dollars, including loans to our allies. Some of the other nations, notably Great Britain, spent even more than this, and the total cost of the war to all the nations engaged has been estimated at over two hundred billion dollars. Owing to the depreciation of monetary units during the war period, such estimates must be used with some caution.

But war costs are not confined to the periods of actual warfare. Maintenance of the military and naval establishments, pensions to veterans and their dependents, and interest (properly a war charge, since war has been the cause of practically all the debts of national governments) make up a total which in the comparatively tranquil first decade and a half of the twentieth century accounted for about half of the total peace-time expenditures of the leading nations of the world. It might be supposed that the United States, with her traditions of peace, her remoteness from warlike neighbors, and her small military establishment, would present a different picture. Actually the expenditures chargeable to past and future wars made up 60 per cent of the total cost of the Federal Government in 1913 and 73 per cent in 1923.

Other national expenditures. Among the other leading expenditures of the United States Government are rivers and harbors and other public works, marine and land transportation, various facilities for national industry, maintenance of justice and security, regulation of private business, sanitary regulations, immigration and naturalization service, and the maintenance of the legislative, judicial, executive, and administrative branches of the government.

State and local expenditures. Education stands first among the expenditures of American state and local governments. Next come highways, care of dependents and defectives, maintenance of justice and security, protection against disease, and interest on debt. The United States Census gives the following statistics of expenditures of the states and the larger cities:

GENERAL DEPARTMENTAL EXPENSES OF STATES AND CITIES, 1923

	States	*146 Cities with Population Over 30,000*
General government	$ 77,749,847	107,514,531
Protection to person and property	52,750,383	244,721,688
Development and conservation of natural resources	46,621,424	. . . . .
Health and sanitation	22,384,439	120,343,465
Highways	108,743,222	106,089,396
Charities, hospitals, and corrections	172,187,600	74,904,045
Schools	347,916,201	444,474,058
Libraries	1,827,306	15,024,687
Recreation	2,604,567	39,521,285
Miscellaneous	63,168,793	55,463,283
Total	895,953,782	1,208,056,438

The growth of public expenditures. The constant growth of public expenditures is one of the outstanding phenomena of human history. Starting as far back as reliable statistics will take us, this universal tendency appears. It applies to all types and grades of government: monarchies and republics, peaceful nations like Switzerland as well as militaristic Germany, imperial governments, national governments, states, and municipalities. The increase in government expenditures before the nineteenth century is on the whole more apparent than real, the great advances in the absolute figures being substantially offset by such factors as diminishing reliance on goods and services directly rendered, growth of population, increase in territory, increased wealth of the people, and depreciation of the monetary standard. But since the beginning of the nineteenth century there can be no doubt that the growth in the cost of government has been both real and substantial.

The limits of our space will require us to confine our further investigation of this topic to the United States, and for the greater part of the nineteenth century the paucity and unreliability of the statistics of state and local expenditure restrict us to study of the expenditures of the Federal Government alone. The figures shown in the following table are significant:[1]

[1] The figures for total expenditures are from *Annual Report of the Secretary of the Treasury on the State of the Finances*, 1925, pp. 486–89.

ORDINARY EXPENDITURES OF THE UNITED STATES GOVERNMENT

Year	*Total Expenditures*	*Per Capita Expenditures*
1792	$ 5,079,000	$ 1.28
1800	10,786,000	2.03
1810	8,156,000	1.13
1820	18,260,000	1.89
1830	15,143,000	1.18
1840	24,317,000	1.42
1850	39,543,000	1.71
1860	63,130,000	2.01
1870	309,653,000	8.03
1880	267,642,000	5.34
1890	318,040,000	5.05
1900	520,860,000	6.85
1910	693,617,000	7.54
1915	760,586,000	7.66
1919	18,952,141,000	180.49
1920	6,141,745,000	58.10
1923	2,464,168,000	21.71

For the first seventy years of our national existence, the expenditures of the Federal Government were on a modest scale, the increase from ten millions in 1800 to 63 millions in 1860 being no more than sufficient to keep pace with the growth of the country's population, as is shown by the figures for per capita expenditure. The figures for 1870 show the effect of the Civil War cost, and the later figures give evidence of the universal truth that war not only adds a war-time burden of extraordinary costs but sets a new scale of peace-time expenditure from which no people has ever been able to go back to the pre-war status. Other influences contributed further to the growth of public expenditures, so that from the Civil War to the World War the cost of the Federal Government averaged more than four times as much as before the Civil War, after making full allowance for the growth of population.

The gigantic costs of the World War reached their climax in 1919. Since then expenditures have declined to about two and a half billions a year, about twenty-two dollars per capita, but it is evident that the nation has once more set a new standard of public expenditure which makes the record of our first century and a quarter appear modest indeed.

Total cost of government in the United States. For the more recent years we have data enabling us to bring state and local expenditures into the picture of the growth of government costs in the United States. The following table is the result of researches of the National Industrial Conference Board.[1]

TOTAL GOVERNMENTAL EXPENDITURES, UNITED STATES

	Total (millions)			*Per Capita*		
	1890	*1913*	*1923*	*1890*	*1913*	*1923*
Federal	$291	$ 692	$ 3,459	$4.61	$ 7.17	$31.26
State	77	383	1,450	1.22	3.97	13.10
Local	487	1,844	5,136	7.72	19.10	46.41
Total	$855	$2,919	$10,045	$13.56	$30.24	$90.77
Nat'l income	$12,082	$34,400	$67,000	$191.61	$356.43	$605.44
Ratio of total expenditures to nat'l income	7.1%	8.5%	15.0%	7.1%	8.5%	15.0%

It will be noted that division of the total cost of government between our three grades of government has not greatly changed in the past generation. The Federal Government accounts for about one third of the total, the states for ten to fifteen per cent, and the local governments for something more than half. This is significant as showing that, in spite of the powerful influence of war in causing expansion of federal expenses, the costs of state and local government have not fallen behind. The climax of total government cost was in 1919. Since then the Federal Government has been exercising economy and reducing expenses. But in the meantime the expenditures of the state and local governments have continued to rise, and it is evident that from now on increases in state and local costs will more than offset reductions in the cost of the Federal Government, so that the total cost of government is again on the increase.

In 1890 seven per cent of the income of the American people was devoted to defraying the cost of government. In 1923, the per-

[1] *Tax Burdens and Public Expenditures*, 1925, page 34. Slight differences between these figures for Federal expenses and those contained in the table on a previous page are due to differences in the methods of compilation, of no importance for our present purpose.

centage was fifteen. The profound significance of this development should be apparent to all. As time goes on a continually greater part of the income of the people is being taken and employed by government. In return valuable services are being rendered by government. Whether this distribution of the national income is giving the maximum of utility to the people and, if so, how far the proportion devoted to government may continue to increase without a decrease in utility are questions to which the student may well give serious thought. They raise problems in economics, in political science, and in social philosophy which, though most inviting, must be regarded outside the field of this elementary treatise.

The cost of government services. All that government expends represents income that has first to be produced by the people, who are then compelled to yield up a part of their income (about one seventh in the United States) in order that the government may exist and perform its functions. This is the price which the people must pay in order that they may receive the various services of government which were passed in review in an earlier chapter.[1] This juxtaposition of government service and government cost is seldom consciously made by the average citizen, who inclines to look with complacency upon government services and government expenditures, to grumble more or less bitterly over the taxes he has to pay, and to fail to put the two together and so to realize that the people cannot "have their cake and eat it too," that whatever services are obtained by the people from government must be paid for by the people out of income which otherwise they might expend for other enjoyments. Government thus becomes a powerful factor in the distribution and consumption of wealth, and the subject of public finance attains a practical importance which few persons fully realize.

Government revenues. To meet the heavy expenditures necessitated by the performance of its functions, the modern government must have an equally large income. In modern times there are just four important sources from which government income is

[1] Chapter X.

to be obtained and which now demand our study.[1] These are (1) industrial earnings (prices), (2) taxes, (3) fees and assessments, and (4) loans.

Industrial earnings. When a government engages in a business enterprise, such as the United States Post Office or the water supply system of a city, it sells its product (goods or service) at a definite price to the consuming public much as any private business concern does. The citizen may take the offered service or not as he wishes, and if he does not take it he makes no payment. Payment is voluntary and in proportion to the amount of the service received. For example the government offers, through the post office, to transmit letters and parcels. If the citizen chooses to take advantage of this service he pays the stated price; otherwise he pays nothing to the post office. In these respects government industry is similar to private business.

The finances of government industry. In contrast with the private entrepreneur, the government's purpose in its industrial undertakings is ordinarily, not the making of a profit, but the furnishing to the people of a service which it is believed can be better provided by the government than by private enterprise. Thus the post office exists, not to make profits for the government, but to afford the people facilities for the widespread communication of messages and information and the cheap and easy transmission of merchandise in a manner which it is believed could not be obtained from private agencies.

In general governments seek to make their industrial enterprises pay their way; prices are fixed so that the receipts shall just about equal the expenditures in each particular industry. There are, it is true, many exceptions. Sometimes a government reserves for itself a monopoly of a certain business and runs the industry frankly for the sake of making a profit for the public treasury. The French tobacco monopoly is an example. This result is scarcely possible

[1] For the sake of completeness and to avoid possible misunderstanding it may be well to mention certain minor sources of government income, whose investigation the limits of this book will preclude. The most important of these are: gifts, forfeiture, reversion, escheat, indemnity, confiscation, fines, and penalties.

unless the government enjoys a monopoly position; otherwise the competition of private business concerns will prevent the government industry securing a profit. On the other hand some government industries do not secure revenues sufficient to meet their expenses, thus showing a deficit which has to be made up out of the general treasury. The United States Post Office is regularly in this state. However, in spite of exceptions, the general rule of public industries is that the income obtained tends, roughly at least, to be equal to the expense incurred. In any event each government industry yields industrial earnings which are earmarked to meet the expenses of that particular industry.

Financing the primary functions of government. If now we transfer our attention from its industrial undertakings to the primary functions of government, we shall find some fundamental differences in their financing. The primary functions, such as defense, justice and security, protection against disease, education, etc., are undertaken by the government in order to benefit the people as a whole, without the purpose of giving particular benefit to particular individuals. For example the state and local governments of the United States provide public schools and universities, at great expense, not in order to benefit those persons who want an education for themselves or their children, but because it is believed that a prosperous and enduring democracy is not possible unless there is widespread education among the people. Education is provided therefore to benefit the whole people, not for the sake of those particular individuals who attend the public schools. Likewise the army and navy exist in order to render the general service of defense to the whole people, not to give particular benefits to particular individuals.

In all these cases therefore no attempt is made to pay the cost of the service by charging a price and so collecting from each citizen in proportion to the benefit he has received. The government does not sell education as it sells postal service. Education is offered freely to all. A may have one child in the public school, while B has half a dozen. No account whatever is taken of these facts in determining the respective amounts which A and B shall pay to-

ward the support of the schools. Furthermore no citizen is able to avoid support of the public school by declining to avail himself of its services. C may send all his children to private schools and there pay for their education, and D may have no children at all. Neither will find it possible to cite these facts as reasons why he should not contribute to the cost of the public schools. The burden of furnishing the money necessary to pay the cost of public education, as of the other primary functions of government, is distributed among the people according to certain rules, which we shall presently study, but there is no necessary relation between the amount of any person's contribution and the service received by him, nor may any person avoid contributing by declining to avail himself of the government's service. Contribution toward the cost of the primary functions of government is compulsory.

It follows that no attempt is made in the case of the primary functions of government to obtain in connection with each function a revenue earmarked to meet the expenditures of that particular function. As a general rule, little or no revenue comes in to the government in connection with the performance of any particular function, and the bulk of the revenue is received without any reference to the particular functions of government for which it is to provide.

Taxes. The foregoing analysis of the purpose of the government's primary functions and the financial contrast between such functions and the government's industrial undertakings has introduced us to the principal characteristics of taxation, namely: (1) compulsion, (2) devotion to the general welfare, and (3) apportionment without reference to individual benefit. It will be found helpful to have a precise definition of a tax; the following will satisfactorily serve our purpose: *A tax is a compulsory payment by a person to the government, destined to defray the cost of government services performed for the common benefit, and paid without reference to individual benefit.* It is upon taxes that modern governments principally rely for obtaining the money necessary to defray the cost of performing their primary functions,

and taxes make up the bulk of the income of most modern governments.

Fees. The primary functions of government are performed, as has been shown, for the common benefit. Nevertheless it often happens that, in performing its primary functions, the government incidentally and unavoidably gives special benefits to particular individuals. Consider for example the granting of patents for inventions. The purpose of this service of government is not to give a special benefit to the inventor; if such were its purpose it would be a case of favoritism, repugnant to the spirit of democratic government, and would not be tolerated. The real purpose of the patent is to promote the public welfare by encouraging the discovery and disclosure of useful new devices and by offering inducement to capitalists and entrepreneurs to develop and market the new devices. It is for this reason that the inventor is given for a time a privileged position of monopoly with respect to his discovery. Nevertheless this is a material benefit conferred upon the inventor, and for that reason the government imposes upon him a special charge, justified by and, somewhat roughly it must be confessed, related to the particular benefit which he is presumed to have received. Such a payment is called a fee.

Fees arise also from somewhat different circumstances. There are certain occupations which cannot be left to the free and unrestricted operation of all persons without danger to the public interests. Such are the businesses of banking, insurance, taxicab operation, peddling, etc. Other examples were cited in Chapter X, where there was also discussion of the reasons for government regulation of such businesses. Now the performance of this regulatory function necessitates government expenditure, and the person who chooses to enter such an occupation is thereby causing expense to the government. The government therefore charges such person a fee, covering part or all of the cost imposed upon it by his engaging in the regulated occupation. Fees are required not only for engaging in certain lines of business, as insurance, peddling, etc., but they are also imposed upon those who engage in any activity over which some government regulation has been found necessary.

One cannot drive an automobile, or own a dog, or get married, without a government license, and in each case he is charged a fee to cover the cost which his act has imposed upon the government.

We now have in mind the nature of the fee, which may be defined in precise terms as follows: *A fee is a payment by a person to the government on account either of a special benefit received from the government or a special cost imposed upon the government in connection with a government service performed for the common benefit.*

Special assessment. Suppose that the city authorities decide to pave a certain street. There will result a material advantage to the owners of land fronting that street on account of the inevitable rise in the value of all such land. The same rise in land values follows the introduction of a sewer system, street curbs, sidewalks, and other improvements. The city government undertakes a new street pavement, not from any desire to benefit particularly the owners of the land, but because the interests of all the people of the city will be promoted by a better surface on that street. Here then is a perfect occasion for a fee and, in America at least, a fee is usually charged, which is none the less a fee in spite of its common designation as a special assessment. This is the definition: *A special assessment is a particular kind of fee, imposed when the special benefit consists in an enhancement of the value of land.* The amount of the special assessment may be anything up to the total addition to the value of the land caused by the government improvement. The special assessment furnishes a most interesting example of the fee, since there is here usually an attempt to relate the amount of the payment to the value of the special benefit conferred much more closely than in the case of fees in general.

Taxes, fees, and industrial earnings compared. Fees are similar to taxes in that they are compulsory. The owner of land whose value is increased cannot decline to accept the benefit and refuse to pay the special assessment. Once the improvement is determined upon, he has no choice in the matter. Likewise the patent fee, the state insurance fee, the motor vehicle operator's fee, etc., are compulsory payments. This is notwithstanding the unquestioned fact that any one might be avoided by keeping out of the occupation

in question. Since the alternative of disbanding the insurance company, declining to take out a patent, giving up one's automobile, or avoiding marriage, in order to escape payment of the purely incidental and usually trivial fee, is quite fanciful, we are justified in regarding the fee as a compulsory payment, like the tax.

On the other hand fees resemble the prices in which the government receives its industrial earnings in that they are paid on account of particular benefits received from, or particular costs imposed upon, the government. Fees thus stand in a midway position between industrial earnings and taxes, and the characteristics of the three sources of revenue may be conveniently represented by a tabular statement such as the following:

Prices	*Fees*	*Taxes*
Voluntary	.....	.....
For particular benefit	For particular benefit	.....
.....	Compulsory	Compulsory
.....	.....	For common benefit

Names may be misleading. In distinguishing prices, fees, and taxes, we must have regard to their real nature and not be misled by mere names. For example a fee is never greater than the value of the special benefit received or the special government cost occasioned by the fee-payer. If this limit is exceeded the payment is, at least to the extent of the excess, a tax. This is usually the case with the automobile operator's " license fee " collected by the states of the United States. The income thus obtained is far in excess of the cost of regulating the operation of motor vehicles and yields a very considerable revenue available for paying the other costs of government. Even when the income is devoted specifically to the repair and maintenance of the highways, the payment is hardly entitled to classification with fees, since the relation between the amount of the payment and the benefit received from using the roads or the injury caused to the roads is so very remote. The motor operator's license fee, so called, is therefore more correctly a tax. On the other hand consider the tax upon gasoline, which has come into recent prominence in the United States. This tax is imposed by the state at so many cents per gallon, and the pro-

ceeds are usually devoted entirely to the maintenance and repair of the highways. When the facts are thus, this so-called "tax" is really a fee. It is a payment which may be justified either on account of the benefit received from the public highways or the cost imposed by wear and tear upon the highways, and it is the most accurate means yet discovered for apportioning the size of the payment to the amount of such benefit or cost. For another example, we may note that the so-called "dog tax" is really a fee.

Relation between expenditure and income. We have divided government functions into two classes: (1) the primary functions, and (2) government industries. This may serve likewise as a classification of government expenditures, except that for that purpose we must add: (3) non-functional expenditures, to take account of expenditures which are not assignable to any particular function of government. The principal items in this class are interest and amortization of debt, and there may also be items for refunds and losses, trust funds, etc., and finally a miscellaneous or catch-all group. The expenditures occasioned by government industries are in the main met by the industrial earnings which arise from the operation of the respective enterprises themselves. The expenditures for the primary functions and for non-functional purposes, together with any deficits resulting from the operation of the government industries, are met normally by taxes, fees, and assessments, and a number of minor forms of income,[1] plus any surpluses which may have come in the way of industrial earnings. By far the greater part of the total income of modern governments comes from taxation.

The total income from the sources named may equal, exceed, or fall short of, the total expenditures. In case of excess, the government has a surplus; in the opposite case it has a deficit. In the latter case only one resource is available; namely, borrowing, which is the fourth item in our classification of government income.

Borrowing to meet a current deficit. A government borrows in order to counteract a deficit, realized or anticipated, and for no other reason. The circumstances which cause the deficit and so

[1] See a partial list in the early part of this chapter.

compel the government to borrow are however various and will now require our attention.

In the absence of extraordinary events which may arise to affect either the expenditures or the income of a government, there is normally an approach toward agreement between the totals on the two sides of the account. Even in the best regulated governments, however, an exact balance is seldom to be looked for. Moderate deficits and surpluses occur from time to time, and, while in the long run they tend to offset each other, the periods of deficit require borrowing to make both ends meet. Such borrowing in order to meet a deficit in the ordinary running expenses is common to all governments. Especially in the United States, where our budgetary procedure and financial administration are far from perfect, Federal, state, and local governments are frequent borrowers for this purpose.

Emergency borrowing: war. A great national emergency, such as a war, is certain to cause a heavy deficit, since no nation can meet the extraordinary expenditures of a great war out of its ordinary sources of revenue, however heroically it may expand its tax system and other resources. Moreover the magnitude of war costs is seldom foreseen by the people and their leaders at the war's beginning, realization of the extraordinary contributions which will be required comes slowly, and the war is always well along, sometimes approaching its close, before the maximum financial effort of which the people are capable is exerted. Often a weak or unintelligent leadership purposely avoids calling upon the people for heavy tax contributions. For these reasons war normally causes large government loans. The American Revolution was almost entirely financed by loans. The cost of our Civil War was about three billion dollars, of which four fifths was borrowed, the balance being met mainly by taxation. The most gigantic loans in the history of public finance were occasioned by the World War. The United States borrowed about twenty-four billion dollars or three fourths of the cost imposed by the war. Great Britain expended forty-two billion dollars on account of the war, obtaining thirty-six billions by means of borrowing. Germany, under the guidance of

a weak and mistaken leadership, borrowed almost the entire amount of her war expenditure. The public debts of practically all nations, even before the World War, were mainly due to war borrowing, while today there is virtually no exception to this generalization. War is the greatest of all causes of borrowing by national governments.

Local governments and members of federal governments like the states of the United States do not have to finance wars, and they are seldom confronted with other serious emergencies. Emergency borrowing is rare with them.

Borrowing for investment or public works. When a government engages in an industrial enterprise of any magnitude it usually finds it necessary to borrow the capital required. When the Prussian Government decided to take over the railway system, it borrowed the money to purchase the existing railroads from their private owners. The United States Government borrowed in order to construct the Panama Canal. The state of New York has recently borrowed large sums to enlarge and improve the Erie Canal. American cities occasionally borrow in order to establish water works systems and other municipal industries. Governments also borrow, not infrequently, the money necessary to pay the cost of buildings and other public works needed for the performance of their primary functions. States and towns frequently borrow to build roads and bridges. Towns and cities borrow to build school houses and other public buildings. Numerous other examples might be added to this list.

The nature of government debts. When the government undertakes to borrow, it must offer terms which will be attractive to the private owners of capital just as any individual or corporation borrower must do. The government has no position of special favor; it can call upon no power of compulsion as in the case of taxes and fees. It is true that a government might conceivably require its citizens to lend to it, and indeed the measures employed by some of the nations engaged in the World War to urge popular subscriptions to their war loans approached at times rather closely to this. But it is generally recognized that real compulsory bor-

rowing partakes of the nature of taxation more than of borrowing. If money is to be taken from the citizens by force, why make any promise to repay? By compulsory borrowing a government would almost certainly impair its credit and the market standing of its public debt, thereby weakening or destroying the real borrowing power which must always be its main reliance in time of emergency. Governments therefore generally choose to enter the loan market like any other borrower, asking no favors except such as may accrue from the magnitude of the operation or the appeal to patriotism. That the latter is an important consideration in time of war, enabling the government to borrow at lower rates of interest and on otherwise more favorable terms and in greater amounts, is not to be doubted. The appeal to patriotism was worked to the limit by all countries in the popular loan campaigns of the World War. But after all, the borrowing government cannot lean too heavily upon any such advantages but must in the long run offer terms which, as a cold business proposition, will induce the owners of capital to lend.

The security back of public debts. Government debts differ from ordinary private obligations in that there is usually no pledge of security to guarantee payment. The exception to this statement relates to weak governments with doubtful credit which, when borrowing from foreign capitalists, are sometimes required to pledge certain public property or certain sources of revenue. Quite commonly the customs administration is turned over to an agent of the foreign creditors, who collects the revenue and, after taking out the agreed sum for interest and amortization of the debt and the expenses incidental to his fiscal administration, turns the balance over to the debtor government. Such pledges are however the exception, not the rule, and in the case of a strong nation no pledge whatever is ordinarily given beyond the word of the government. In case of default of the ordinary public debt, the creditor has no recourse except either to bring suit in the debtor government's own courts, which is not much help if the government has deliberately chosen not to pay, or if a foreigner to seek to persuade his own government to take diplomatic or, in the last

resort, military action against the defaulting government. The history of public finance is replete with cases illustrating all of these principles.

The high standing of government obligations. It may appear strange, in the face of the foregoing, that the unsecured debts of sovereign powers generally enjoy such high standing. As a matter of fact government obligations, especially those of strong national governments, which are almost invariably unsecured, are ordinarily held in high favor by the investing public. This is shown by the fact that, except when its credit is weakened by a war of uncertain outcome or by some other emergency which threatens its national position, a government can generally borrow at lower rates than are paid by other borrowers, and government obligations are generally quoted on the markets at higher prices in proportion to their yield than any other investments. This is true, only in lesser degree, of the bonds of many states and cities. Government bonds stand in the very front rank of " gilt-edged " securities.

The high standing of government obligations rests on certain characteristics of government. In the first place, the life of a government is not limited as is the life of an individual, and it is considered more enduring even than the lives of most corporations. More fundamentally, the government's ability to meet its obligations, both interest and principal, is not dependent upon the success of any particular business venture, even though the loan may have been made to obtain capital for an industrial enterprise. Most borrowing by individuals and corporations is for investment in business. If the business fails the borrowers may have to default on the interest or on part or the whole of the principal. The government's ability to meet its obligations rests upon the taxing power, which may always be exerted to obtain the necessary money provided the debt is not too large. The absence of security is counteracted by a most scrupulous regard for its financial obligations, by which most nations seek to protect their credit. It is well realized that only thus can a nation preserve the indispensable resource of its credit for time of war or other national emergency. Political attacks upon the holders of government bonds have

threatened the inviolability of government credit more than once; the financial history of the United States furnishes some very instructive examples. But the determination of the responsible national leaders, backed by the honest common sense of the majority of the people, has generally been sufficient to maintain the public faith. In public finance no less than private, indeed even more so, "honesty is the best policy."

The national debt of the United States. The history of government debts is a most interesting chapter in public finance, but one which the limits of this book will not permit us to enter beyond a brief glance at the history of the public debt of the United States. As has been stated above, the War of the Revolution was financed principally by loans, and when the new national government of the United States started operations in 1789 it had a heritage of 52 millions of public debt, a sum which appears trivial enough to us of the present generation, accustomed to speak most casually in terms of billions, but which the people of that day regarded as a staggering burden upon the shoulders of the new government.

Nevertheless one of the first acts of the government, under the far-sighted leadership of Alexander Hamilton, was the assumption and addition to the national debt of the debts of the several states incurred in the prosecution of the Revolutionary War. This brought the debt up to 76 millions in 1794. American public opinion, from the very beginning, has looked with disfavor upon a large national debt and has always favored rapid payment. By 1812 the debt of the United States had thus been reduced to 45 millions. The War of 1812, likewise financed principally by loans, brought the national debt up to 128 millions. Then commenced a period of rapid debt payment, culminating in 1836 in the complete extinguishing of the national debt except for a few thousand dollars, not paid because the holders had not presented the bonds and for whose payment the Treasury had the cash in hand. This represents a unique performance in the records of public finance — a national debt completely paid off.

The country was not long out of debt. Deficiencies in current revenues and the Mexican War restored the debt, which just before

the Civil War in 1860 amounted to 65 million dollars. Reference has been made to the borrowing of this war, which provided four fifths of the three billion dollars which that tremendous struggle cost the United States; on August 31, 1865, the national debt stood at $2,845,000,000. Once more the nation set about the task of debt payment and, by a performance only less remarkable than its previous record, had brought the debt in 1893 down to 839 millions, a reduction of more than two thirds. Since then there have been numerous debt operations, increases as well as reductions, the principal causes of new borrowing being the resumption of specie payment of the Greenbacks in 1879, the currency disorders of 1893–1896, the financing of the Spanish War, and the digging of the Panama Canal. By 1916, the year before the United States entered the World War, the national debt had made a net increase of about fifty per cent, amounting to about a billion and a quarter. The World War was financed to the extent of three fourths of its cost to the United States by borrowing, principally by the issue of the Liberty loans, Treasury notes, certificates of indebtedness, and war savings certificates. The total borrowing on account of this war makes everything preceding pale into relative insignificance; on June 30, 1919, the national debt amounted to $25,482,000,000. And now again the traditional American policy of debt reduction is being vigorously pursued. Large appropriations for payment of the debt are included in the regular annual budgets, and the debt has already, January 31, 1926, been reduced to $20,283,000,000, a reduction of more than one fifth of the debt in six years.

The World War debts. We cannot pursue this topic into the debt histories of other nations. Suffice it to call attention to the overwhelming effect of the World War upon the indebtedness of most of the leading nations, referring the reader to the figures given in a previous section of this chapter. One final observation of great significance must be made. The unprecedented borrowing of the World War has done much to threaten the high standing of government obligations, which was the subject of an earlier paragraph of this chapter. As a result of budget deficits and reckless

issue of paper money, Russia, Germany, and certain other great nations have virtually repudiated their earlier public debts. Other nations, not in quite so bad a way as this, have nevertheless seen their obligations sink far below par and have suffered serious loss of credit; among these are France and Italy. Great Britain and the United States are the two conspicuous examples of nations which came through the struggle without serious impairment of national credit. It remains to be seen whether, after time has permitted recovery from the most disastrous financial effects of the War, our paragraph upon the high standing of government bonds may have to be qualified by limiting it to only a few great nations.

Government surplus. If a deficit in the government's finances is to be regarded with anxiety, a surplus is generally worse. This rather startling statement may require defense. In private finance a surplus is of all things the most fervently to be desired. Faced with a glittering array of possibilities for its disposal, the idea that it may become an embarrassment is the last to enter the mind of the fortunate one. But we are here confronted with one of the several important differences between public and private finance. Suppose that a sovereign government (the only kind of government which, with proper financial management, should ever have a formidable surplus) finds itself, on account of some unexpected accretion of revenue or simply as the result of bad budgeting, in the possession of a large surplus. What, let us ask, shall be done with it? Simply to store money in the vaults of the treasury is not permissible, since the withdrawal of so material a part of the country's money from circulation will have those disastrous effects upon prices, business, and private contracts which the reader now fully understands. Investment naturally occurs; this is the obvious solution of the problem in private finance. But the spectacle of a democratic government announcing itself as a would-be purchaser of the stocks and bonds of business corporations or going quietly into the market and purchasing securities presents a prospect of favoritism, corruption, and public loss which will be readily visualized by anyone with only the slightest knowledge of political science. May not the surplus be used to reduce the public debt?

Here is indeed the one safe outlet, and so long as surplus revenues are devoted to retirement of debt no serious consequences will ensue and the result will be wholly favorable. Unfortunately the presence of surplus money offers so great a temptation to the politicians in control of the government's finances that it is not certain that the safety-valve of debt payment will always be permitted to operate to the full extent.

The United States has had some interesting experiences with surplus revenues, which illustrate these principles. Beginning immediately after the close of the War of 1812, the growing customs revenue and the fortunate sales of public lands brought receipts in excess of the government's needs, so that for twenty years there was an almost uninterrupted condition of surplus revenue. Until 1836 the public debt furnished the ready means of disposal, and no serious embarrassment was felt. But in that year the debt had been extinguished; and the surplus continued. Congress and the people were in a quandary, and, after a long wrangle, the entertaining details of which we must pass by, it was decided to divide the surplus among the states. An outright gift appearing likely to be held unconstitutional, the money was transferred to the states as a loan with the tacit understanding that its repayment would never be demanded. The problem was thus passed on to the states, which disposed of the money in a variety of ways but generally without any great permanent profit. Among other consequences, the surplus and its disposition were partly responsible for the disastrous panic of 1837, which incidentally solved the problem by turning the government's surplus into an embarrassing deficit.

The next period of surplus culminated in the decade 1880–1890. In August, 1888, the government's surplus was equal to one fifth of the country's entire monetary circulation. Now again the public debt offered a convenient outlet, and we have already noted the extraordinary record of debt payment which came to an end in this decade. However it did not prove feasible to use the entire surplus in this way, and the tempting hoard made political fingers itch. Congress, incited by the Administration, proceeded to "blow in" the surplus in a riot of extravagant expenditures for pensions,

river and harbor improvements, and public buildings, setting a standard from which the taxpayers have been suffering ever since.

There have been occasions of surplus revenue since, notably in 1907 just before the panic of that year, and these surpluses have always been troublesome. Enough evidence has been presented to sustain the general principle that in public finance a surplus is more to be feared than a deficit and presents so embarrassing a problem that few democratic governments are able to meet the test.

The budget. Under sovereign constitutional governments all taxes and other revenues and all expenditures must be authorized by law. This requires that at certain stated intervals the law-making body, be it Parliament, Congress, state legislature, or other legislative authority, must take account of the state of the public finances, determine what services shall be performed by the government, estimate the probable cost of such services, and provide taxes and other income to meet the anticipated expenditures. This calls attention to still another difference between public and private finance. The ordinary procedure in government finance involves first the determination of expenditure for a certain period in the future and thereafter the exaction from the people of revenues sufficient to pay the bills. Private persons seldom enjoy this privilege of making their incomes fit their desired expenditures. With them the order is reversed; the future income must first be calculated as accurately as may be, after which the objects upon which that sum is to be expended may be chosen; and expenditures in excess of income actually received present a serious situation, not easily to be adjusted by a loan, as is the practice of governments. The individual is held strictly to the obligation "to cut his coat to fit his cloth." Of course it is not to be inferred that governments are wholly independent of this rule. Even government finance is controlled in the last analysis by the national resources, which set a final limit to expenditure. But there is in ordinary times a wide margin before this limit is reached, so that the normal financial plans of government are not much affected by it. Witness the extraordinary expansion of public income from

taxes and loans which every strong nation finds possible in time of war or other national emergency. Government always has a degree of control over its income which the ordinary private person does not possess.

In determining its fiscal policy, the legislature will find it helpful to take account of the income and expenditure of the past year or recent years. Such account, together with the estimates and plans for the coming year, is called the *budget.*

The subordinate grades of government, counties, cities, towns, school districts, etc., also have use for the budget. Although they have no sovereign power, they have been delegated a considerable authority over their own finances, under which the local legislative body — board of alderman, city council, town selectmen, or whatever it may be — determines the expenditures and at least the rates of taxation for the district.

Budget procedure has been brought to a high degree of perfection by certain governments, the most notable being the British Parliament. In the United States the budget has been generally more honored in the breach than in the observance by all grades of government. Its efficient development has been hindered in the Federal Government by our constitutional separation between the executive and legislative arms of government and our bicameral Congress, principles which have been generally carried down into the political organization of our states; and there have been other contributing causes of our poor record. Only since the World War has the Federal Government developed any semblance of a real budgetary procedure, and, barring a few notable exceptions, the states and municipalities have been no less backward.

CHAPTER XLV

TAXATION: GENERAL PRINCIPLES

How is the tax burden distributed? Almost the entire revenue required to meet the cost of the government's primary functions comes in the form of taxes, which are exacted from the people with little or no regard to particular benefits received. It is our present task to learn how the burden of taxation is distributed — what rule determines how much each person must pay for the general services of government.

Equal distribution. It may perhaps serve to clarify the first steps of our investigation if we stop for a moment to consider the rule of equal distribution. This will be recognized as the rule which usually determines the contributions of the several members for the support of the general services of a club; the "dues" of the country club, the athletic association, the debating society, the ordinary social club are normally the same for all the members, and no other rule is apt to be considered seriously. Conceivably such a rule might be adopted for distribution of the tax burden; let us see how it would work in the United States. Of the ten billions of total government expenditures in the year 1923, seven and three quarters billions were raised by taxation. With a population of about 110 millions, the share of each person — man, woman, and child — would be seventy dollars.[1] At this rate a family of five (father, mother, and three children) would have an annual tax bill of $350, no account of course being taken of the size of the family income. If the man's wages or salary and other income were $1,000 his tax bill would be more than one third of his income. Another family of five, with an income of $10,000, in paying the same amount

[1] The exact figures, as calculated by the National Industrial Conference Board (*Tax Burdens and Public Expenditures*, 1925, page 12) are: Total taxes, federal, state, and local, $7,716,000; per capita taxes, federal, state, and local, $69.72.

would pay only about one thirtieth of their income, while the very wealthy would find their tax burdens quite insignificant. To the very poor the burden would be absolutely crushing, if indeed they were able to pay at all.

Now, quite apart from any question of justice, it is evident that this rule would be unworkable, since there are large numbers of the people who could not possibly pay at the rate of seventy dollars a year for each taxpayer and all those dependent upon him. And, whatever one's standard of justice may be, it is safe to say that virtually everyone would condemn this plan as flagrantly unjust, placing as it does so crushing a burden upon the poor and taking so little from the rich who might easily pay more. The important conclusion is that public opinion tolerates, even demands, a plan of tax distribution which makes some persons pay more than others. Having cut loose from the base of equal distribution, what shall be our rule of apportionment?

Benefit received. There have been those who believed that the costs of government should and could be distributed on the basis of the benefits received by the individual members of the group. Of course this is the rule in business dealings, and we have learned that to it government industries are no exception. Even in the case of the primary functions, there are times when an individual benefit is evident, making possible the collection of a fee, though often the fee, while justified by the presence of a special benefit, can scarcely be said to have its amount determined accurately in relation to the benefit received. As to the great bulk of government services, direct individual benefits are seldom discernible and practically never capable of measurement. How should we apportion to each citizen of the United States the benefit received by him from the activities of the army and the navy? What has the Supreme Court of the United States done for him? What part of the President's salary should he pay on the basis of benefit received? How should we determine the contribution of the citizen whose life has been saved by a naval and military demonstration against a lawless foreign state? And if we could determine the value of such a service, how could the government collect, except

from the very wealthy? The government must exert its every power to protect all its citizens rich and poor alike; it cannot follow the rule of serving only those who are able to pay what the service is worth. The rule of payment according to benefit would indicate heavy payments as due from the paupers in the almshouses and the inmates of state institutions, since they receive everything from the government. Clearly this basis is not workable as a general rule for the apportionment of the tax burden. At the same time it must not be discarded altogether. Just as it serves well as the basis for fees and assessments, so it has a certain bearing at least upon the obligation to pay taxes, if not upon the amount of the individual contribution. This will come up again in our study of particular taxes.

Cost of service. It has also been proposed that taxes be apportioned upon the basis of the cost of the service rendered each person by the government. This principle is very like the principle of benefit received, and its criticism follows similar lines. Here also we have a basis which has been availed of, when it was applicable, to indicate the obligation to pay a fee and, roughly at least, the amount of the fee. Here also we meet the same obstacles when it is proposed to apply the rule generally. What part of the cost of the United States navy is imposed upon it by a peaceful and law-abiding Illinois coal miner who has never been outside his home state? And if it cost the navy and marines half a million dollars to save a citizen who fell into the clutches of a tribe of savages, shall he be made to pay the cost? And if he be a poor man, shall he be protected only to the extent of his ability to reimburse the government for the cost? Here again we conclude that this cannot ever become the general rule for tax distribution, while yet it may serve a useful purpose in certain cases which are on the border line between taxes and fees.

Ability to pay. The almost instinctive rejection of the plan of equal distribution of the tax burden is undoubtedly based upon the common feeling that a system of taxation which inflicts a crushing burden upon the poor while imposing virtually no sacrifice at all upon the wealthy is unjust. It is felt that those who have much

should pay more than those who have little, that there should be at least some approach to equality of sacrifice, that people should pay taxes in proportion to their ability to pay. There is no doubt a lack of philosophical precision in these popular notions, but they are widely held and with sufficient clearness to indicate another rule for tax apportionment; *i.e.*, the principle of ability to pay, which may be paraphrased accurately enough as the principle of equal sacrifice. This rule evidently puts the distribution of taxation upon a basis quite distinct from those that have been previously mentioned, and it thereby avoids many of the weaknesses of these preceding rules. It does not lead to grotesque contrasts, as does the rule of equality. It arrives at no absurdities in connection with paupers and defectives, since it permits of certain citizens paying nothing. It does not require impossible calculations of the individual benefit or the individual cost of services which are for general, not individual, benefit; nor does it compel the government to limit its services by the separate amounts that can be collected from the individual citizens.

It is not our purpose to undertake a philosophical analysis of the principle of payment according to ability. Its broad meaning should be clear enough for our present purpose. As so understood, the ability principle has on the whole commended itself to modern students of taxation and to modern public opinion as the basic principle of tax apportionment, to be supplemented by the principles of benefit and cost so far as they may be applicable. It must never be forgotten however that the complex practical problems of taxation are not to be solved by the mechanical application of any set of general principles. Modifications of general rules and resort to other principles will be found necessary to meet the exigencies of particular kinds of taxes. The foregoing is a statement of our general conclusions as to the basis of tax apportionment, and we may now proceed to the study of certain practical questions concerned with the application of the principle of ability, during which some further light may possibly be thrown upon the meaning of the principle itself.

How is ability to pay measured? The collection of taxes being a

practical financial operation, there is necessity of a financial yardstick with which to measure taxpaying ability. For all practical purposes, just two measures are available: (1) the possession of wealth (or property), (2) the receipt of income. We may fairly assume that to the student who has reached this point in the study of economics the appropriateness of these two measures will be evident.

As to both of these measures it is generally the net amount, rather than the gross, that more accurately measures taxpaying ability. A and B may each own wealth of a value of $50,000. If A has debts of $30,000, representing property rights of others against him and his wealth, while B has no such claims against him, it is evident that the taxpaying abilities of the two, so far as indicated by wealth or property, are not the same, and that A's ability is to be measured by his net possessions of $20,000. Speaking strictly and as we have previously defined the terms, it is not wealth, but property, that is the real measure. In brief the measure is the net property, being the difference between all the taxpayer's possessions, including property rights against the wealth and persons of others, and the property rights of others against him and his wealth. Likewise in using the income measure, it is net income; *i.e.*, the difference between gross income and the expenses necessary to secure such income, that is the real measure of taxable ability. Further and more detailed application of these principles will be made later, especially when we come to the study of property and income taxes.

Proportional taxation. In the application of the property and income yardsticks to the measurement of taxable ability, there arises at once the question whether the relation between wealth or income and taxable ability is one of simple proportion or is something different. Suppose that A has a net income valued at $2,000, while B's is worth $4,000, and C's, $40,000. Does this indicate that B should pay twice as much as A, and that C should pay ten times as much as B and twenty times as much as A? If so we have proportional taxation, to be accomplished by a uniform *ad valorem* rate. Let the rate be four per cent. Then we have this result, in

which the tax payments are in direct proportion to the respective net incomes:

Person	*Net Income*	*Tax Rate*	*Amount of Tax*
A	$ 2,000	4%	$ 80
B	4,000	4%	160
C	40,000	4%	1,600

Progressive taxation. If, instead of making the tax payments directly proportional to the base (either property or income), the tax payments are made to increase at a rate faster than the increase in the base, we have progressive taxation. For example, suppose a scheme of income tax rates such as this:

On the first $1,000 of net income			1%
On the part of the net income:			
over $ 1,000, up to and including		$ 2,000	2%
over 2,000, up to and including		3,000	3%
over 3,000, up to and including		5,000	4%
over 5,000, up to and including		10,000	5%
over 10,000, up to and including		20,000	6%
over 20,000, up to and including		40,000	7%
over 40,000, up to and including		60,000	8%
over 60,000, up to and including		80,000	9%
over 80,000, up to and including		100,000	10%
over 100,000,			11%

Now A, with a net income of $2,000, will pay a tax of 1% of his first $1,000 plus 2% of his second $1,000; *i.e.*, $10 + $20 = $30. Similarly B's tax would be calculated thus:

1% of the first	$1,000,	$ 10
2% of the second	1,000,	20
3% of the third	1,000,	30
4% of the fourth	1,000,	40
Total tax on	$4,000,	$100

C's tax would amount to $2,390. These results may be tabulated thus:

Person	*Net Income*	*Tax Rate*	*Amount of Tax*	*Ratio of Tax to Net Income*
A	$ 2,000	1%–2%	$ 30	1½%
B	4,000	1%–4%	100	2½%
C	40,000	1%–7%	2,390	5.975%

Under such a set of rates, B's tax would be more than three times that of A, though his income is only double A's. C's income is twenty times A's, but his tax is practically eighty times as great. This is the result of progressive taxation and is accomplished, be it noted, by giving up the uniform rate in favor of a rate which increases with increases in the base.

It is probably not possible to prove scientifically that progressive taxation is required by the principle of ability. For that we should need a calculus for the measurement of sacrifice which thus far does not seem to have been discovered. However there is little doubt that the consensus of opinion, expert and general, inclines to the support of progression. As the matter is commonly put, $100 taken from a man with an income of $1,000 (10%) means a good deal more in the way of sacrifice than $1,000 taken from an income of $10,000, or $10,000 taken from $100,000. Progressive taxation is a comparatively modern idea, and its development in the past generation or two has been rapid. We may confidently count upon the continued and increasing presence of this principle in modern tax legislation.

If there is no scientific proof of the correctness of progressive taxation, still less can there be any scientific determination of the exact scale of progressive rates required to apply the ability principle. In actual legislation the rates of progressive taxes are determined principally by rule of thumb, guided by various principles and interests, harmonious or conflicting. This is a matter to which we shall later return.

Regressive taxation. It is possible to apply a principle opposite to progression, by fixing rates which decrease as the tax base increases. This is called regressive taxation. There are few if any who defend this rule as an accurate application of the ability principle, and it practically never appears consciously embodied in tax legislation. However many taxes are regressive in their actual results, as will appear in our later study.

The tax-free minimum. There are those in any community from whom it is useless to attempt to collect taxes. Of course all those who are not even capable of supporting themselves are not able to

make any contribution toward the support of government, and there are also those just on the line between self-support and dependency to whom any tax burden at all may prove the last crushing straw. If nevertheless taxes are actually collected from such, there must be corresponding return in the form of charity, either public or private. We are reminded of the English laborer who comforted his fellow, grumbling over the heavy tax on tobacco, by reminding him: "Well, if they don't tax our tobacco, how can they pay our unemployment doles?"

Tax theory, if not always tax practice, therefore assumes that a certain minimum of subsistence shall be exempt. This softens the burden of proportional rates and intensifies the favorable discrimination of progressive rates to the poorer taxpayers. If, in the examples just used to illustrate proportional and progressive taxation, there had been an exempt minimum of $1,000, the taxes due from the three citizens would have been as follows:

			Proportional Tax		*Progressive Tax*	
Person	*Net Income*	*Taxable Income*	*Tax*	*Ratio to Net Income*	*Tax*	*Ratio to Net Income*
A	$ 2,000	$ 1,000	$ 40	2%	$ 10	½%
B	4,000	3,000	120	3%	60	1½%
C	40,000	39,000	1,560	3.9%	2,320	5.8%

A comparison of this table with those above shows that the exemption of a moderate amount (1) makes a nominally proportional tax really progressive and (2) gives a very important reduction to those taxpayers having small incomes while being negligible in the case of the wealthy.

Other measures of tax obligation: consumption taxes. Property and income are clearly enough the most appropriate measures of taxpaying ability. That the principle of ability is not alone resorted to in the framing of actual tax systems is evident from the presence of other very important yardsticks for measuring tax obligation besides income and property. There are for example the "consumption" taxes, such as the stamp taxes levied by the United States Government upon tobacco and playing cards, the tax on theatre tickets, etc., under which the amount of each individual's payment is in proportion to his use of the taxed commodity

or service. Now it is evident that a person's consumption of any particular article, while doubtless in some degree related to his income and therefore to his taxpaying ability, is at best an extremely crude and inaccurate measure of such ability. The same is true of a general consumption tax, such as a general sales tax or a "spendings tax." Such taxes fail to correspond to ability because they take no account of that part of the taxpayer's income which is saved. Suppose A and B are neighbors with about the same needs and the same scale of expenditures, each spending around $3,000 a year, although A's income is $6,000 while B's is $3,000. A's taxable ability is double that of B, if not more, and is clearly not reached to its full extent by a tax on consumption or spending, which takes the same amount from him as from B. The relation between tax payment and ability is still more remote when the consumption tax is, as is usual, levied upon a selected list of commodities or services instead of upon all expenditures.

Since the consumption taxes do not reach that part of income which is saved and since saving is in general the privilege of the well-to-do and the rich rather than of the poor, it follows that consumption taxes have generally a regressive effect. The same result follows from the fact that the articles selected for levy by most governments are those of wide general use, which therefore tend to demand a larger part of the total incomes of the poor than of the rich. The regressive character of taxes on consumption and expenditure further demonstrate the departure of such taxes from the simon pure ability principle.

Business taxes and poll taxes. Taxes imposed on account of engaging in business and measured more or less accurately by the amount of business done or the amount of business profits received are quite common. Such taxes are well known in Europe. In America they are to be found in the business license taxes of many of the states, particularly in the South, in the old liquor licenses imposed by nearly every town and city before prohibition, and in most of the special taxes on corporations. When imposed in addition to, not in lieu of, the regular taxes upon the owner on account of the property employed in the business or the income derived from

it, business taxes can with difficulty be related to the principle of ability, but must generally be frankly acknowledged as applications of the benefit or the cost principle. The assumption is that government gives special protection to property employed in business and special services to those persons, natural or corporate, so engaged, or that the cost of government is especially increased on account of business activities of the people.

The poll tax, being usually a uniform contribution from all individuals, has no relation whatever to ability, but is related to the benefit principle, on the somewhat dubious theory that protection of life is granted to all and is of equal value to all.

All taxation cannot be apportioned according to ability. The fact that consumption taxes, spendings taxes, business taxes, poll taxes, and possibly certain other forms of taxation are not in full harmony with the principle of ability is not to be taken as an *a priori* condemnation of all such taxes, but is rather an indication of the truth of our earlier statement that no one principle can be relied upon to shape all the taxes which any government finds it necessary to impose. It must further be remembered that we cannot pick out a single tax from a tax system and criticize it as though it stood alone. Final judgment must be upon the tax system as a whole. Thus it is possible to combine certain taxes measured according to benefit or cost with other taxes in such manner that the whole tax system may have reasonable conformity to the principle of ability. Likewise it is possible to combine regressive consumption taxes with progressive income taxes, inheritance taxes, etc., so as to obtain a tax system which may be either proportional or progressive in its general effect.

Escape from taxation. No one likes to pay taxes. It is only human nature that every taxpayer should use all lawful means to make his taxes as light as possible and that some taxpayers should resort to measures that are not lawful. Certain modern taxes, such as the property tax, the income tax, and the inheritance tax, are so involved with complexities that it is often difficult for the taxpayer to know just what his legal obligation is. For example there can be no doubt that, under the United States income tax, many per-

sons are paying more than their lawful taxes, through misunderstanding of the law or mere carelessness. Under such circumstances the use by some taxpayers at least of every lawful means, including the employment of tax experts and lawyers, in order to avoid paying more than is legally due, is inevitable and justifiable. Again a person may escape a heavy tax upon his property by selling his real estate and moving to another town where the tax rate is lower, or a man may avoid the inheritance tax by giving his property to his intended heirs during his lifetime. Finally it is often possible to shift the burden of a tax to other persons through increasing the price charged for some commodity or service. Put in general terms, it may be possible to escape a tax burden that might otherwise have to be borne, (1) by exercising care not to pay more than is legally due, (2) by putting one's self in such position that the tax law is not applicable, or (3) by using the power which a tax may give to raise the price of something sold.

Now all of these expedients may be perfectly legal and, if so, it can raise no ethical controversy to assert that they are entirely justifiable. For the first the legislature itself is principally responsible, through enacting tax laws too complicated or obscure for the understanding of the ordinary intelligence. In the second case, if it appears that taxpayers are finding it easy so to adjust their residences or businesses or possessions or other circumstances as to put themselves outside the application of a tax, thus defeating the purpose of the law, that again must be due to defects in the tax itself, for which the lawmaking body alone is responsible. The case of shifting the burden to others may indicate either failure of the law to place the final burden upon those whom the legislature intended to reach or simply the result foreseen and intended by the legislature. If the first, then the miscarriage is not to be ascribed to any dereliction, legal or moral, on the part of the taxpayer, but simply to lack of understanding of economic laws on the part of the lawmaking body. As to the second, we have the simple and common case of indirect taxation, which will be examined presently, but which requires no further comment in this connection.

In popular discussion there often appears a certain failure to

distinguish between such legal and proper escape from the burdens of taxation and the evasion of taxes legally due. However much this popular confusion may be excused by the obscurity of tax laws and the complexities of the whole modern tax situation, the student must always keep the distinction in mind. Escape from taxation may be lawfully accomplished in a variety of ways; the term "evasion" should be reserved for violations of the tax laws.

Tax evasion. When tax evasion occurs in sporadic instances the situation is similar to any other case of violation of law and is of no particular concern to the economist or the tax student. When, as is sometimes the case, there is wholesale evasion, the responsibility must be either upon the legislature for enacting a law incapable of enforcement or upon the administrative personnel. To ascribe it to a low state of public morality does not face the problem. The legislature has to deal with human nature as it is and is responsible for enacting laws which can be enforced and for providing the necessary machinery for efficient administration. We shall find instructive applications of these various principles when we come to the study of the different kinds of taxes.

The shifting of taxes. Taxes are *shifted* when the person upon whom the government has imposed the tax finds that the presence of the tax has given him the power, which he did not before possess, to raise the price of some commodity or service which he has for sale, thereby reimbursing himself in whole or in part for his tax payment and to that extent passing the burden on to someone else. An example familiar to everyone is the United States internal revenue tax upon the manufacturers of tobacco products. A comparatively heavy tax is thus imposed, but the manufacturers, who are the immediate taxpayers, regularly pass the burden of the tax on to the consumers of tobacco through an increase of the price. The problem of shifting is simply one of applying the economic laws of price to the circumstances of each of the several kinds of taxes. It is better studied therefore in connection with the investigation of the respective taxes. For the present we need only to acquire a general notion of the subject. The degree of shifting which is possible varies, according to the nature of the tax, all the way from

those whose burden is entirely or almost entirely shifted, such as the tobacco tax just cited, to those whose shifting is impossible, such as a tax upon monopoly profits or upon pure economic rent. Shifting and *incidence* (meaning the final resting place of the tax burden) must never be overlooked when considering the problem of the distribution of the tax burden or when criticizing any particular tax. Obviously the person who finally bears the burden is more to be considered than he who pays only to shift the burden to others.

There are cases where the government, in collecting a tax, does not regard the one who pays as the taxpayer at all but merely as the agent who makes payment for the taxpayer. In such case the one who pays is of course expected to reimburse himself entirely by collecting from the real taxpayer. It may be questioned whether this should be called tax shifting at all, but, however that may be, all such cases are fully within the scope of the principles which have been formulated.

Direct and indirect taxes. Taxes are usually put into two classes according as the burden is or is not normally shifted by the one from whom the government collects the tax. Direct taxes are those whose burden remains on the whole upon those from whom the tax is originally collected. Indirect taxes are those whose burden is normally shifted. Since the possibility of shifting may be present only partially and in various degrees, it is not always possible to draw a sharp exclusive line between the two classes of taxes. However the characteristics of the more important kinds of taxes are so clearly marked that there is usually little difficulty in placing any given tax in its proper class. This distinction, be it noted, is an economic one. It does not agree precisely with the legal definition of a direct tax, as laid down by the Supreme Court of the United States, about which a long and interesting controversy has raged in United States tax history, with which we shall presently become better acquainted. There is no occasion here for dispute. We have simply to deal with two different concepts, the economic and the legal, and must be on our guard only against confusion from the use of the same technical terms in two different senses.

The purpose of taxation. Of course the primary purpose of taxation is the raising of money whereby the government may pay the cost of performing its primary functions. In the discussion thus far we have ignored any other purpose. The time has now arrived when we must recognize that taxation may be imposed for other purposes and must examine the economic bearing of the more important of these purposes. The taxing power is among the most powerful and far-reaching of the attributes of sovereignty. Even when applied only for the purpose of securing government income its indirect economic effects may be, indeed certainly will be, very great. When consciously used for the accomplishment of other ends its power can scarcely be exaggerated. Taxation may be used to impoverish one individual and enrich another. It may foster or kill whole industries and destroy the prosperity of entire regions. It may divert the course of industry and trade, affect the movement and location of population, interfere with the distribution and consumption of wealth. With such a power in its hands, we shall not be surprised to find the government occasionally using its power for the accomplishment of ulterior ends, not to speak of the numerous programs of social reform through taxation which are constantly being urged upon the government and the public.

The most familiar illustration is the protective tariff. The tariff, or customs, is a tax, imposed originally and primarily for the sake of revenue. At various times and places it has been used for the purpose of checking the import of certain foreign products in order to relieve home producers of foreign competition and so to build up the home industry. This purpose has been so very prominent in the United States during most of our history that there may be a question whether protection is not really the primary purpose of the American tariff. Another example is the ten per cent tax imposed by the United States upon the notes of state banks. This tax was introduced with the sole purpose of legislating the state bank circulation out of existence. Revenue was not expected and has not appeared, since the tax has been successful in accomplishing its real purpose. Taxes, as we have learned, are imposed upon the notes issued by the Federal Reserve Banks under certain circum-

stances for the sake of restricting the amount issued, and the same device is employed by other countries. In the days before prohibition in the United States heavy license taxes were generally imposed upon the saloons, partly for the sake of revenue but also for the purpose of restricting and regulating the retail liquor establishments. Among suggestions of the use of taxation for "reform" are the proposal to limit or abolish roadside billboards by heavy taxes, the demand that inheritance taxes be used to produce a more equal distribution of wealth, and the time-honored "single tax" of Henry George.

In general it may be said that the use of a tax for another purpose tends to impair its effectiveness as a revenue machine. This is obviously true to the limit of such a tax as the prohibitory levy upon state bank notes, which produces no revenue at all, and it applies especially to the protective tariff. This subject will be pursued further in connection with our study of particular taxes, especially the tariff.

Some technical terms. A good many popular misunderstandings of tax matters come from confusion as to certain technical terms. Before concluding this general discussion, we need to be sure that these terms are clearly defined and understood. The person from whom the government collects the tax in the first instance is called the *immediate payer*. He is to be distinguished from the *subject*, or person who bears the ultimate burden of the tax after the process of shifting has done its work. Of course the immediate payer and the subject are one and the same in the case of a tax which is not shifted. The term *taxpayer* is somewhat loosely used. It should preferably be applied to the subject when he is a different person from the immediate payer; certainly this should be the usage when the immediate payer is merely acting as agent for someone else.

The *object* or *base* of a tax is the thing (person, commodity, or event) which is used as the measure of the amount of the tax obligation of each person. Usually the object is expressed in terms of money, being the value of taxable property, the amount of income, the value of goods manufactured or sold, the value of goods

imported, etc., though a tax may have for its object a physical quantity of goods, as a tax of one cent a gallon on gasoline sold, or an act or occupation, as the automobile operator's tax, or even the person of the taxpayer, as the poll tax. The object is to be distinguished from the *source* of a tax, this being the fund from which the taxpayer obtains the money for payment of the tax whether directly to the government or through another person who is the immediate payer.

All of these terms are well illustrated by the United States internal revenue tax on cigars. The immediate payer is the cigar manufacturer; the subject, properly called also the taxpayer, is the one who purchases cigars for use; the object is the cigars; and the source is whatever fund the taxpayer draws upon for the purchase of cigars, since the purchase price includes the tax.

A tax is said to be *ad valorem* when its rate is expressed as a certain percentage or other part of the value of the object or base, for example a tax of fifteen mills upon property, of five per cent upon incomes, etc. A *specific* tax is one whose rate is expressed as so much per physical unit of the base, as the tax of one cent per gallon upon gasoline sold, the poll tax of two dollars per poll, etc. Ad valorem and specific duties are both used in connection with the customs tariff.

CHAPTER XLVI

THE GENERAL PROPERTY TAX

From the earliest colonial settlements down to the present day the chief financial reliance of the American colonies and states and of their minor subdivisions has been upon the general property tax. This statement still holds true, in spite of a recent decline in the relative importance of this tax. It is fitting therefore that we choose this as the first form of taxation for our special study.

Evolution of the general property tax. The possession of property furnishes so obvious a basis for the distribution of the costs of government or so obviously invites the exactions of the necessitous ruler, that it is not surprising to find that the property tax is among the oldest of taxes; history shows us numerous examples in ancient and medieval times. In those days and indeed for some centuries thereafter there did not exist the bewildering variety of different kinds of wealth and property rights which surround us of the present. Wealth consisted principally of land and buildings; next in importance came domestic animals and agricultural tools and equipment; household furniture, clothing, and other forms of personal property were scarce and crude and, except in the homes of the kings and nobles, of little value. The earliest property taxes were generally imposed upon land and buildings. Gradually, as other kinds of wealth came into being and importance, they were added to the list of taxables, until finally the struggle to make the tax law keep pace with the growing complexity of human possessions was abandoned, and the tax law was made to apply to every kind of property not specifically exempted, instead of to those kinds only that were specifically mentioned as taxable. When this point was reached the taxes upon property had become the *general property tax*. It later appeared however that the dis-

covery and assessment of the hundred and one kinds of property which were now taxable, such as household furniture, clothing, jewelry, books, and the intangible rights to wealth in the physical possession of others, was a superhuman task, far different from the simple assessment of land and buildings under the original property tax. The evolution then turned back and retraced its steps. By successive additions to the exempt list one class of property after another was relieved of the tax, until finally the property tax returned to about where it started as a tax upon land and buildings and not much else. This is a very brief and summary account of an evolution which throughout most of Europe had run its course by the end of the seventeenth century. Only in the cantons of Switzerland and to a lesser degree in the German states has the general property tax persisted down to the present. It is there no longer the principal source of revenue, and it is imposed at very moderate rates.

In America. In the American colonies and states the general property tax has gone through the first half of this evolutionary process, except where the original settlers brought a more or less developed general property tax with them and so skipped the earlier stages of the process. Having reached the point of its highest development as an inclusive tax upon all kinds of property except those for some reason especially exempted, the general property tax in America has not submitted to the second half of the evolutionary cycle which in Europe decreed its retrogression to a real estate tax. We have clung tenaciously in most of our states to the idea of the general property tax, with consequences which have been most extraordinary, not to say disastrous.

Typical organization and operation. Before proceeding further in this direction, let us try to form a picture of the practical operation of the general property tax in the typical state. Being a subject of state law, the property tax is not carried out in exactly the same way in any two of the forty-eight states. However the general features of the tax present enough similarity to make it possible to construct a composite picture which will represent, not exactly the situation in any particular state, but rather the typical

procedure which, with appropriate modifications, will conform closely enough to the situation in most of the states. The general property tax is provided by statute for the purpose of enabling the states and the counties, cities, towns or townships, and other local jurisdictions to secure the revenues necessary to their needs. The law determines what classes of property shall be taxable and what shall be exempt and prescribes a mass of administrative details. About the only legislative discretion left to the subordinate bodies is the power to fix the tax rates which shall be applied within their respective jurisdictions. The selection of assessors, collectors, and other administrative officers and the actual work of administration however is almost wholly in the hands of either the towns or the counties, although there is a certain power of supervision and control lodged in the state tax commissioner or similar officer or board, the extent of which control however varies greatly among the states.

Assessment. The first step in the operation of the general property tax is the *assessment*, which means the discovery and valuation of all the taxable property of each person liable to the tax, as of a certain date each year specified in the statute. In most of the states assessment is the function of the towns; in other states it is performed by the counties. In this description we shall assume that the town has the matter in charge. The assessor must make out a "tax list" for each taxpayer, showing in more or less detail the description and value of all his taxable property. In some states the law requires the taxpayer to make out his own list and hand it to the assessor. In other states no list is required of the taxpayer, while in still others the taxpayer may be required to give certain information but not all, or he may have the option of filing a list or of having the list made by the assessor with an addition to the amount assessed as a penalty for failure to file the list. In every case however the final assessment is the work of the assessor, who is responsible for determining the total value of property to be assessed to each taxpayer. The law generally gives the assessor ample power to demand all necessary information from the taxpayer or from other witnesses.

Appeal, revision, and equalization. In each town there is a legally constituted body, known as the "board of relief" or by similar title, whose function is to hear and settle appeals from any taxpayer who may feel that his property has been wrongly assessed. When this board has finished its work the tax lists, as corrected, are combined to form the "town list," which contains at the least a list of the names of all taxpayers in the town with the total value of the property assessed against each.

Each town then reports the total amount of its tax list to a county board, commonly called the "board of equalization." The principal duty of this board is to determine after investigation whether the assessment in the several towns in its jurisdiction has been performed according to law, particularly whether the total value of property as assessed is equal to the full true value in each town. The board may arbitrarily increase or decrease the valuation of any town in order to make it equal to the full value, or at least to make the ratio between assessed and true value the same in all the towns. The reason for this equalization is that, as will presently appear, the county tax is to be apportioned among the towns in proportion to their respective tax lists, and any town which undervalues its taxable property or, more precisely, undervalues it in greater degree than the average of all the towns is thereby escaping a part of its share of the county tax and correspondingly burdening the other towns. Sometimes, though not usually, the county board of equalization may have authority to revise individual assessments.

Having corrected the lists of such towns as in its opinion require revision, the county board makes up its county tax list, which is a list of all the towns with the total assessed value of the property in each town. The total for the whole county is then reported to the "state board of equalization" or corresponding board or official. This board performs with respect to the several counties the same process of equalization as has just been performed by each county board with respect to its towns. The purpose also is similar; namely, to prevent evasion of part of its share of the state tax by any county through undervaluation of its taxable prop-

erty. The final result, after the state board has made its revisions, is the state tax list, showing the total assessed value of property in the entire state as distributed among the several counties.

Apportionment; the tax rate. The next step is the *apportionment* of the tax. Each jurisdiction decides what amount of income it must obtain from the general property tax; this amount divided by the total of the taxable property within its jurisdiction gives the *tax rate* for that particular jurisdiction. The tax rate applicable to any given taxpayer will be the sum of the rates of all the jurisdictions by which he is taxable. To make this concrete, let us assume an example. Suppose that the work of assessment in a certain state has been completed, showing a tax list for the whole state of $1,000,000,000, which is distributed among the several counties as follows: County A, $200,000,000, County B, $100,000,000, and the rest in the other counties. Consider now County A, whose tax list of $200,000,000 is the sum of the lists of Town *a*, $50,000,000, Town *b*, $5,000,000, and the several other towns making up the remaining $145,000,000. Let us take account also of one other town, Town *c*, located in County B and having a tax list of $10,000,000, which is of course included in the $100,000,000 list of that county. Now let it be assumed that the State decides to raise $2,000,000 from the general property tax; the necessary tax rate will be calculated by dividing this amount by the state tax list, thus: $\frac{2{,}000{,}000}{1{,}000{,}000{,}000} = .002$, which in ordinary tax parlance would be called "2 mills on the dollar" or simply "2 mills." In exactly this way each jurisdiction decides how much it must raise from the general property tax and then calculates its tax rate. Let the following be the assumed tax lists and amounts to be raised, with the resulting tax rates:

Jurisdiction	*Tax List*	*Amount to be Raised*	*Tax Rate*
State	$1,000,000,000	$2,000,000	2 mills
County A	200,000,000	200,000	1 mill
Town *a*	50,000,000	500,000	10 mills
Town *b*	5,000,000	100,000	20 mills
County B	100,000,000	200,000	2 mills
Town *c*	10,000,000	150,000	15 mills

We can now get back to the individual taxpayer. Suppose three individuals: M, living in Town *a*, N, in Town *b*, and P, in Town *c*, each having property assessed at $10,000. From M there will be collected the taxes of the State, County A, and Town *a*, to whom he must pay respectively 2 mills, 1 mill, and 10 mills upon each dollar of his assessed valuation. His total tax will then be at the rate of 13 mills and will amount to $130. N will pay at the rate of 23 mills, being the sum of the rates of the State, County A, and Town *b;* his tax will be $230. C has a rate of 19 mills, made up of the rates of the State, County B, and Town *c;* his tax will be $190.

State, county, and town (township or city) are the three most important jurisdictions having a hand in the general property tax. But there are various other political bodies which, while having little or nothing to do with the administration of the tax, are permitted by law to collect a tax from the taxpayers within their geographical limits. Thus a borough is generally a part of a town which, being more thickly settled and having need for special town services not performed for the whole town, is permitted to vote for such special services to be paid for by a borough tax at a sufficient rate, as determined by the people of the borough or their proper officers. If for example the individual M in our previous illustration lived in such a borough within the Town *a* and the borough tax rate were 7 mills, his total tax would be at the rate of 20 mills and would amount to $200, whereas another resident of the town living outside the borough would pay only the 13 mill rate. The same situation arises in the case of cities, whenever the city is not coterminous with the town. The school districts are the most important and numerous of the minor taxing jurisdictions; they very commonly have independent authority to determine their school revenues and their district school tax rates. There are also fire districts, street lighting districts, road districts, and other examples of limited groups of the people who are organized for the sake of obtaining special government services with authority to tax themselves by adding their own rates to the general property tax.

In general each person is taxable upon all his taxable property and pays a tax at a rate which is the sum of the rates set by all the political units which have tax jurisdiction over him or his property.

Collection. The final act is the *collection* of the tax. Each taxpayer receives a tax bill, or formal notice of the amount of his total assessment, the tax rate, and the amount of the tax due from him. In some states each town and county and the state has its own collector, who sends out bills and collects the tax due. The taxpayer may thus receive two, three, or more separate tax bills and make as many separate payments. In other states one bill is sent out and the entire tax collected by the town or county collector, who then distributes its proper share to each jurisdiction concerned.

Modifications. The reader should again be reminded that the foregoing description is intended to be typical rather than descriptive of any particular state, and that an almost infinite number of modifications is to be found. Equalization is in some states not performed by both county and state boards. There are certain states in which the state government has given up all share in the proceeds of the general property tax and so has little or no interest in equalization. The process of apportionment is not always carried out in all the refinement here described, and in at least one state an entirely different method is employed. There are states in which certain of the tax rates are fixed by statute in advance of assessment, so that the assessment determines the amount of the yield. Such modifications in detail, of which these are only a few examples, are numerous enough but need not concern us here.

Position of the general property tax. The tax whose practical machinery we have thus examined at some length is the main source of revenue of the states and local bodies of the United States. In the year 1923 according to the United States census [1] it produced for all the states 353 million dollars, 30 per cent of their total revenue receipts of 1,247 millions. As regards the towns and other

[1] *Financial Statistics of States*, 1923.

local bodies exact figures are not available, but it may safely be stated that the yield of this tax is in the neighborhood of three billion dollars a year and contributes between 90 and 95 per cent of all the revenues of the local governments of the United States.[1] As we have seen, the total tax bill of the American people (to all grades of government, federal, state, and local) in 1923 was nearly seven and three quarters billion dollars. The general property tax must have produced more than three and one third billions, exceeding the amount of all taxes levied by the Federal Government and being not very far short of half the total national tax bill in that year. There can be no question of the predominant position of the general property tax in American public finance or of the tremendous consequences of its efficient or inefficient operation. Our next problem is to learn how this great tax engine is actually working.

Critical analysis of operation. In view of the heavy reliance of the American state and local governments upon the general property tax, it will doubtless be a disappointment to the student who does not already know the facts to be told that the general property tax is working so badly that it may fairly be called a failure; in other words, that Europe was wiser than America in abandoning the idea of the general property tax two hundred years ago. That this statement is anything but an exaggeration will appear clearly enough after we have examined critically the way the tax operates in America.

Definitions. This is an appropriate place to define certain technical terms commonly used in property tax discussions. *Real estate* means in general land (in its broadest sense), all buildings and other structures permanently attached to the land, and such equipment, machinery, etc., as is so definitely attached to a building that it could not be removed to another location without serious impairment of its value or use, such as elevators, fire escapes, shafting for machinery, etc.

[1] See the following reports of the National Industrial Conference Board: *Taxation and National Income*, 1922; *Tax Burdens and Exemptions*, 1923; and *Tax Burdens and Public Expenditures*, 1925; also reports of the United States Bureau of the Census upon financial statistics of states and of cities.

Personal estate or *personal property* or *personalty* means all possessions except real estate. Personal property is divided into two classes: (1) *tangible* personal property, being that which has a physical substance, such as machinery, household furniture, jewelry, etc., and (2) *intangible* personal property, which includes all other personalty; that is, abstract rights against wealth in the physical possession of others or against other persons, such as promissory notes, corporation stocks, bank deposits, etc. It will be noted that the word "property" is here somewhat loosely used in conflict with the economic definition of that term. Strictly speaking, intangible personal property alone deserves the title of "property"; the others — real estate and tangible personalty — are "wealth." In most discussions of taxation however the term "property" is broadened to include all physical wealth as well as abstract property rights, and this usage is so well-nigh universal that it is scarcely to be avoided. If we occasionally slip into such usage, no harm will be done so long as the sense in which the word is used is clearly understood.

Assessment: discovery of taxable property. The crucial point in the practical operation of any tax is the assessment. The general property tax requires that there be entered opposite the name of each person a list of his taxable possessions (which includes nearly everything that he owns) with a statement of the true market value of each item and of the total. This implies (1) the discovery of all such property by the assessor in the face of at least the passive resistance of the taxpayer, and (2) the process of valuation. As to the first, the task is easy enough in the case of land and buildings and the more bulky and less movable kinds of personal wealth, such as farm and factory machinery, farm animals, etc. But the task of locating other kinds of tangible personalty, such as jewelry, is far from easy. Even such property as household furniture, rugs, books, pictures, clothing, etc., can be discovered only by a personal search of the taxpayer's home, and although the tax assessor generally has the legal right or even obligation of such search, he is loath to resort to it in the face of the natural repugnance of the taxpayer. House-to-house search

for personal property is today the exception, not the rule, in tax assessment. If it were the rule, it is safe to say that a considerable amount of the more valuable and easily movable articles would be spirited away and concealed on tax day.

When we come to intangible property, the difficulty of discovery becomes practical impossibility. Bonds, notes, stocks, bank deposits, and similar classes of property are readily concealed in bank vaults or in the taxpayer's office or home or if necessary may be actually removed into another town or state over tax day. The assessor is powerless, and the best he can do is to accept such information as is volunteered by the more conscientious taxpayers or to make a pure guess at the amount of such property and trust to the taxpayers to protest in case he has overshot the mark. Intangible property such as real estate mortgages, which must be publicly recorded, are discoverable; but this is no help when the mortgaged property is located and the mortgage recorded in another jurisdiction from that of the taxpayer's residence and tax obligation, since the assessors can hardly be expected to search the records of other towns. Moreover many states exempt mortgages upon real estate within the state, thus removing the only ones that could easily be discovered.

Valuation. The discovery of taxable property is evidently a difficult problem; it is easy compared with the task of valuation. The state laws differ as to language, but in general they require that all property be assessed at its true market value, or words to that effect, sometimes specifying that it must be the value at which the property would be sold by a willing seller to a willing buyer, seeking quite properly to exclude the evidence of forced or otherwise unrepresentative sales. Tax legislation thus agrees closely with our economic concept of value.

Now except at the very time of a sale, the value at which anything "would sell" — in short its *value* — is a matter of judgment. It must be determined by an appraisal, which is just what assessment really is. Anyone acquainted with practical affairs has an idea of what is meant by appraisal. He knows that it can be accomplished only by an expert in the particular kind of property

involved who is in possession of all the pertinent facts. Land appraisals are made for business purposes by real estate experts, who first gather carefully all possible evidence and then give ample time to the study of the problem. They are usually assured of the coöperation of the owner or at least of someone practically interested in the matter. The appraisal of a large apartment house or office building or factory is a difficult technical matter, which only the specialized expert can handle. And here also, when we pass from real estate to personalty, we pass from the difficult to the more difficult. The appraisal of the machinery in a cotton mill, of the raw materials, the " goods in process," and the finished products of a brass factory, the stock in trade — " inventory " — of a dry-goods store is each a complicated, time-consuming task for the man especially equipped for that particular appraisal. If it is instead a silk factory or a hardware store, another set of experts will be required. Such appraisals are constantly being made in practical business, for insurance companies, banks, investors, brokers, and for the manufacturers and merchants themselves. To be accomplished with the accuracy demanded for business purposes, they require the careful labors of specialized technical experts, highly trained and experienced.

Theoretical questions. Furthermore there are theoretical problems of valuation which are not solved merely by employing the right experts. How for example shall we assess partly finished goods — " goods in process " — of a manufacturing plant? The manufacturer has his own methods for inventory purposes, based upon more or less arbitrary accounting rules: he may take cost of raw materials, add labor cost, put in something for " overhead," etc. But, except as an accounting fiction, the " value " thus obtained means little; it is to be doubted if it is really value at all. Who would buy the partly finished goods of most factories? Certainly it is questionable whether a " value " thus obtained is the true market value prescribed by the tax law. And if this is not the legal tax value, what in the world is? Another accounting rule is " cost less depreciation." But if there is no one likely to want to buy the article, can we accept the results

of such a "valuation"? What for example should be the method of assessing special tools or patterns made and used in a factory on a particular job which is now completed? These are merely a few samples of some of the theoretical and technical questions which may well puzzle the expert appraiser or the lawyer.

The assessor's problem. Now let us picture for a moment the town assessor as he enters a great factory devoted to the manufacture of clocks and watches, with millions of dollars' worth of land, water power, factory buildings, machinery and equipment, raw materials, goods in process, finished products, not to mention such intangible property as patents, contracts, "good-will," investments of various sorts, and so on indefinitely. An appraisal of this plant and property may have been recently made for the owners or for a syndicate of bankers interested, which took weeks or months of labor by a corps of experts and cost hundreds or thousands of dollars. The assessor, if not the only one in the town, is one of two or three or half a dozen. He is perhaps a farmer or a carpenter; he was never in a watch factory before, never saw or even heard of many of the kinds of wealth before him. This factory is simply one of hundreds or thousands of properties which he must assess in the space of two or three months at most. He cannot take the time even to glance at all the items of wealth here assembled. Yet he is the one who is required by law to put a value upon this complicated assembly of wealth and property, which shall in large measure determine the amount of the contribution of its owners toward the cost of the town, county, and state governments. Is anything more required to demonstrate the impossibility of his task?

Expert assessment the exception. Now it is true that certain states have obtained some degree of centralization of the work of assessment in the office of a state tax commissioner or board, which has enabled them to secure full-time services of engineers and other experts to aid in the work of local assessment. Certain of the larger cities have tax departments of considerable efficiency. Even at best however such tax organizations are powerless to cope with the gigantic problems of assessment of all the property of a

state or a city, and such as they are they are the rare exceptions, not the rule. The picture painted in the preceding paragraph may be accepted as fairly typical of the impossible situation of the ordinary town or county assessor.

Information from the owner. It may be asked: why not compel the owner to give the assessor the advantage of the results of his own appraisal, or inventory, or accounts; or does not the law in fact now give the assessor power to demand such information from the taxpayer? The answer is that, the tax assessor being unable to check the accuracy of such reports, reliance upon them would result in a system of self-assessment. Since even the most expert appraisal is a matter of judgment and since there is the greatest variety in accounting practice, errors and inequalities would be inevitable, even assuming an honest intention on the taxpayer's part to make a truthful return. On the other hand with the strongest of motives and almost unlimited opportunity, the probability of false statements and dishonest collusion is too great to be ignored. American tax practice has never fully accepted the principle of self-assessment, and, while a very considerable amount of information is required of the taxpayer by the law, the ultimate obligation of the assessor to view and value the taxable property is seldom waived. Furthermore the very natural disinclination of the taxpayer to disclose to any outsider the intimate details of his personal possessions or the technical facts of his business is so strong that it has generally resisted successfully the assessor's efforts to secure complete information.

Where assessment is easiest. It must also be pointed out that all assessments are not so complicated and so far above the heads of the assessors. The local assessor in an agricultural town or a small village, himself a farmer or carpenter or blacksmith, gets on after a fashion in the assessment of farm lands, buildings, and equipment, and of the simple homes and personal effects of the village. Even the stock in trade of the country store may be fairly within his ken, and he may not come up against the really difficult problems. But let no one delude himself with the thought that even this simple kind of assessment is as a rule even fairly

well done; the facts are all to the contrary. And it is the exceptional town which does not contain some examples of taxable property that will put to rout the equipment of any ordinary assessor.

The attitude of the taxpayer. With such an impossible task upon him and with human nature as it is, the assessor can hardly expect much help from the taxpayers. There are those indeed, and their number is surprisingly large, who conscientiously try to pay all that the law requires, even though there be little to compel it. Such persons are rewarded for their extreme honesty by bearing a burden of taxation out of all proportion to that borne by the rest of the community. The majority take the attitude of passive resistance: "Let's see you get it if you can." And a few use every means, fair and foul, to evade as much of their tax burden as is possible. So general is the cynical acceptance of this situation that the majority of honest, law-abiding citizens regularly once a year take oath to tax lists which they know perfectly are false on their face, and if any taxpayer was ever prosecuted for this perfectly open perjury we have not yet heard of the case.

Undervaluation. Assessed values are almost everywhere below true values; sometimes by a moderate margin, sometimes so far that they are only a small fraction of the real value. Assessments at 20 per cent, 15 per cent, 10 per cent, of true value or even less, are no curiosities. There are two reasons. First is the natural desire of almost every taxpayer to get his own property in at the lowest possible value, in order that he may gain at the expense of his neighbors. In the second place is the common interest of the whole town in a low scale of assessments, since thereby the town reduces its share of the apportioned county and state taxes. Each taxpayer is therefore glad to have all assessments low and would in addition like to see his own at least a little lower than the general average. Such universal pressure is too much for the assessor, generally owing his position to the votes of the taxpayers, to resist.

How the assessor meets the situation. Under such circumstances as these, we need not be greatly surprised at the result.

The assessor generally does about the best he can. In the first place he has last year's lists, of which he takes full advantage; copying last year's figures will suffice for the majority of cases where he has reason to believe there has been no material change. He will travel over his district, making brief inspections of the outside of new buildings, casting cursory glances at the merchants' shelves, asking a few questions, picking up what gossip he can as to the worldly goods of his fellow townsmen. Using such information as this, reinforced by a good deal of often shrewd guesswork, influenced almost always by the popular satisfaction which will follow when everybody feels that he has been assessed considerably under true value, influenced now and then no doubt by the desire to be kind to a friend and to stand well with those of political and social influence, seldom downright corrupt, the assessor manages somehow to check up on the lists handed in by taxpayers, to make out the lists of those who fail to file them, and finally to turn in a town tax list which will not cause too many appeals to the board of relief, nor be too much censured by the county or state board of equalization, nor cause so much local dissatisfaction as to threaten his reëlection. This description, like those that have preceded it, is intended as a composite picture of the typical procedure in the ordinary American town or county. It must be remembered that there are exceptions, especially in the larger cities, where the work of assessment is highly organized with permanent trained personnel and where the assessment, of real estate at least, attains a certain degree of scientific precision. If the assessment of all kinds of property in all places were as efficient as the assessment of real estate in some places, the property tax might be a success. Unfortunately this is not the case. Real estate is assessed at least moderately well in a few cities only; personal property is nowhere well assessed.

The results of assessment. With all the circumstances as thus pictured, we should hardly expect to find a true and faithful assessment of all taxable property, and the facts will fully justify our surmise. For reasons which are now obvious, the assessment of real estate is more effective than that of personal property. Yet

figures collected by the United States Census, going as far back as 1890, show that the assessed value of real estate all over the country is not more than half the true value legally taxable. These figures are confirmed by investigations by state tax commissions in various states and by numerous tax students. And while on the average real estate gets onto the tax lists at not more than half its true amount and value, this gives no adequate picture of the extreme variation among individual properties, which are assessed all the way from zero (when the property is not on the tax list at all) to 100 per cent of the value or even more. A careful investigation made in 1922 by the Connecticut State Tax Commissioner [1] showed in one small Connecticut town no item of real estate assessed at more than 40 per cent of its true value, while the lowest was 9.2 per cent of true value and the average of all the cases investigated was 24 per cent. In another town, somewhat larger, the percentages of assessed to true value ranged all the way from 41 per cent to 106 per cent and the average for the thirty cases investigated was 76 per cent. Note not only the extraordinary variation among particular assessments but the great discrepancy between the two towns. For the whole state the average ratio of assessed to true value of real estate is probably not over 60 per cent. There is similar evidence from other states and no reason to believe that any state would show conditions much better.

When we examine the results of similar investigations of the assessment of personal property, the figures are either appalling or ridiculous according to one's state of mind. The United States Census, beginning with 1890, finds that less than 20 per cent of the true value of taxable personalty gets on the tax lists. As far back as 1870 a New York state commission estimated the ratio at 15 per cent in New York. There is every reason to believe that the breakdown is becoming more complete from year to year. A special investigation conducted by the Connecticut State Chamber of Commerce in 1916 [2] showed the following re-

[1] *Report of the Tax Commissioner* (of Connecticut), for 1921 and 1922, pp. 6–16.
[2] *Report of the Joint Committee on Taxation and State Finance to the State Chamber of Commerce*, 1917, pp. 12–18.

sults as regards the assessment of personal property upon the farms of the state: only 55 per cent of the true value of cattle was listed for taxation, 17 per cent of the cattle being not assessed at all; of sheep, goats, swine, and poultry less than 24 per cent of the true value was assessed; 41 per cent of the true value of horses and mules was reached; and the assessed value of farm machinery was 6 per cent of its true value. And it must be remembered that these are among the classes of personal property most readily discoverable and most easy to value. Investigations into the assessment of household furniture, books and libraries, musical instruments, watches and jewelry, and other such classes of personal property demonstrate the almost complete escape of such possessions. And finally, when we come to intangible personal property, the value actually assessed is nowhere more than a trifling fraction of the amount of such property in existence and legally taxable. It is no exaggeration to say that the escape of intangible personal property throughout the United States is almost complete and that in most parts of the country the assessment of important classes of tangible personalty is either a ridiculous failure or has been virtually abandoned through breakdown of the administration.

Results of the breakdown of assessment. Of course as certain classes of taxable property escape in whole or in part from the general property tax, either through failure to be listed at all or through undervaluation, the burden upon other forms of wealth which are more effectively reached is correspondingly increased. The states, counties, and towns must have their necessary revenues, and, as the tax base is narrowed by concealment and undervaluation, the tax rates have to rise. Tax rates have been going up steadily all over the United States, virtually since the beginning. Rates of 20 mills, 30 mills, 40 mills are common; 80 and 90 mill rates are not unknown. Such rates mean confiscation except when tempered by undervaluation. Undervaluation is the rule for most kinds of tangible wealth, real and personal, but for many kinds of intangible property undervaluation is not possible. If the assessor discovers a $10,000 mortgage at a good rate of interest

upon adequate security he can scarcely enter it on the tax list at any other value with that figure staring him in the face. Corporation stocks and bonds, regularly quoted in the market reports, can hardly be entered at anything but their true market values. Now a tax at the rate of 2 per cent to 4 per cent upon a 4 per cent bond means anywhere from taking half the value to complete confiscation. This will be resisted by any intelligent taxpayer whose conscience is not quite quixotic and will indeed make the assessor himself hesitate. Intangible property is virtually forced whenever possible to the escape of complete evasion, and we have the vicious circle in perfect operation: the higher the rates the more complete will be the evasion of all property which can possibly escape, and the more property succeeds in escaping the higher will the rates have to be on that which is left. Such property as cannot escape assessment is able to endure prevailing tax rates only because of prevailing undervaluation, and thus low valuations and high tax rates chase each other about another vicious circle. Real estate has thus come to bear the bulk of the tax burden; at least four fifths, probably more, of the property assessed consists of real estate.

Inequality and the legal remedy. With complete escape of certain kinds of property the regular thing, with gross undervaluation the general rule for all property that cannot escape, with tax rates so high as to be utterly unendurable except when accompanied by undervaluation, with variation in individual assessments all the way from complete escape to more than 100 per cent of true value, necessity of demonstrating inevitable and outrageous inequality among the several taxpayers scarcely exists. We can safely let it go at that, and proceed to inquire as to the operation of the legal machinery for relief.

Any taxpayer who feels himself unjustly treated on the assessment of his own property may present his complaint before the local board of review. But, except in the few cases where there is actual assessment at more than the true value, this avails him nothing. If in a town where property is assessed on the average

at 40 per cent of true value, a certain taxpayer is assessed at 80 per cent, it is obvious enough that he is unjustly treated. But all he can prove to the board of relief is that his property, even so, is assessed 20 per cent less than the law requires. There have been a few cases in which the courts have required reduction of a particular assessment to agree with the general practice, but since proof of a general practice usually requires bringing evidence as to the true value of taxable property of every other taxpayer, this remedy is virtually nonexistent. Justice by legal review can be effective only when the millennium of full value assessment shall have been reached.

The attempt to give relief for wholesale inequality through general undervaluation by towns and counties, is equally ineffective. The county board of equalization has usually no possible way of determining the true value of taxable property in each of the towns. The most it can do is to gather evidence as to a few particular assessments, listen to the gossip that is current, and mainly by guesswork raise this town and lower that in the hope of diminishing some of the most glaring inequalities. The review of the state board of equalization is equally ineffective, though there are a few states in which a competent tax commission, armed with authority to revise original assessments in the towns, has been able to accomplish some good.

Theoretical weaknesses of the general property tax. Attention thus far has been focused upon the practical operation of the general property tax and its administrative breakdown, but we must not overlook certain fundamental defects in the very nature of the general property tax idea. In the first place we note a serious confusion between the concepts of wealth and property. If all physical wealth were taxed and in addition all the property rights to wealth, there would be no injustice, since wealth and property rights to wealth are coterminous; there would simply be an unnecessary duplication of taxes where one tax, on either wealth or property at double the rate, would accomplish the same result. But many forms of property have no legal document to represent them and hence do not come within the purview of the tax laws.

The result is that our laws usually attempt to tax *all* forms of physical wealth (barring the special exemptions) and in addition *some* classes of property rights. For example A and B are merchants, each owning a store and fixtures and stock in trade assessed at $100,000. A incorporates his business, the entire capital stock being owned by himself and members of his family. Now if, as some of the states require, he is taxed upon the store and fixtures and goods assessed at $100,000 and also upon $100,000 of capital stock in the corporation, his tax is twice as much as it was before and twice as much as B's tax, with absolutely no difference in taxpaying ability. This is double taxation of the crudest sort. The same thing results when the holder of a mortgage or promissory note or other credit is taxed upon it, while the debtor is taxed upon the entire value of his possessions without deduction of the amount of the debt. Yet permission to deduct debts is the exception in American tax legislation.

The theory of the general property tax discloses another weakness when it is admitted that the possession of property is far from a true measure of taxpaying ability. Without going into an exhaustive discussion, it will be recognized that receipt of income, rather than possession of property, is the real measure of ability to pay taxes. Now it is true that generally and in the long run property is merely the capitalized value of income to be received; but at any given time the relation between current income and value of property may be quite remote. There is much valuable property which is for the time being yielding no income to its owner. Still more serious is the fact that many persons are the recipients of considerable incomes which are not represented by any forms of taxable property. Such are the incomes from wages, salaries, professional fees, etc. Even as early as our colonial days, this defect was seen and a remedy attempted through the addition of the so-called "faculty tax" to the basic general property tax.

Modifications in detail. The reader is now in possession of some notion of the nature of the general property tax, its historical development, its typical organization and operation in the United

States, and the results in those states which have not subjected it to material modifications. The farcical failure of the typical general property tax is a fact that can no longer be questioned; its complete breakdown is rapidly approaching, where it has not either already occurred or been forestalled by modifications which, where sufficient to be effective, have amounted essentially to abandonment of the general property tax idea.

Mention may be made of certain remedies in detail which have served somewhat to improve the operation of the general property tax. In some states the state has withdrawn from participation in the revenue from the tax, thus removing one cause of undervaluation. But if county apportionment continues, the improvement is not great, and there may be actual loss through withdrawal of state supervision and control. The state of Connecticut has solved this problem in the one effective way by apportioning the state tax, not upon the assessed value of property in the towns, but upon the respective amounts raised by the property tax.[1] Since reduced assessments must be accompanied by higher local tax rates in order that the town may secure its own revenue (the product of assessment by tax rate), the town has nothing to gain by undervaluation so far as the state tax upon the town is concerned.

Recognition of the peculiar position of intangible property, which has been pointed out, has led a number of states to provide for the taxation of such property at rates much lower than those applied to tangible wealth. It has been the hope that the taxpayers when granted a reasonable rate would voluntarily list their intangibles, and the experience of several states has shown a material increase in the amount of such property assessed following a reduction in the rate. There is no evidence however to show that any state has by this expedient succeeded in assessing the whole or even the greater part of the intangible property subject to its tax laws, and there is considerable evidence to show that the low-rate remedy gives at best only a slight mitigation of the former evasion.

[1] Connecticut has unfortunately continued to reward undervaluation by certain grants of state moneys to the towns with low tax lists.

Reform of the general property tax. Adequate treatment of the topic indicated by this title can not be compressed within the limits of a general text such as the present. A few suggestions only will be offered. Reform of the *general* property tax is not possible; only by abandoning the idea of a general property tax can relief be expected. The first step is certainly the abandonment of all taxation of intangible personal property. This is indicated both by the theoretical nature of intangible property rights and by the complete and inevitable breakdown in administration. This step has been recommended by numerous tax authorities, including a committee of the National Tax Association. Next in order is the abandonment of the attempt to tax all those classes of tangible personalty where experience has shown the practical impossibility of a real assessment. This would include household furniture, private libraries, musical instruments, clothing and jewelry, tools and machinery, goods and materials of merchants and manufacturers, and various other categories. In brief we may as well admit that these classes of property never have been and never will be properly assessed and give up attempting the impossible. This would leave on the tax roll only real estate and a few forms of bulky, easily discovered, and easily valued personal property. What should be substituted in order to reach the taxpaying ability possibly represented by the classes of tangible property exempted is a question which must be dismissed here with the suggestion that some form of the income tax offers the logical substitute. It can be pointed out here however that, since the classes of property which it is proposed to exempt are the ones which now generally escape and which yield comparatively little revenue, the loss of revenue to be made up from this part of the reform is not great.

We must still take account of the fact that the present assessment even of real estate is far from perfect. Here however the task is not inherently impossible, and there is reasonable hope for improvement, once the assessors are relieved of the obligation to do the impossible and are allowed to concentrate their efforts upon that which can be done. As already suggested, there are certain

cities whose assessment of lands and buildings has attained a high degree of efficiency. Accurate maps of the entire region are made and skilfully used. Complete descriptive data regarding land and buildings are prepared and recorded. Evidence regarding all sales is gathered. Scientific rules for determining values with respect to location and various other standards are used. Even the aeroplane is pressed into service for the quick mapping of extended areas in the suburbs or country districts. A permanent staff of trained assessors, with the necessary clerical and other assistance, is built up. Such an organization is possible in any town or county where the finances are on a large enough scale to warrant the expense, and for the smaller towns a consolidation of the work of assessment in the hands of larger units, as the counties, is possible. While the work of assessing real estate will never be so organized as to produce perfect results, it may still bring about a condition of reasonable equity, infinitely superior to the present chaos.

A certain degree of centralization of administrative authority in the hands of a state officer or board is a great help in the task of assessment. This has been obtained in some states with beneficial results. It makes possible the employment of technical experts whom no one town or county could afford. Statistical data, rules of appraisal, and assistance of various sorts may be furnished the local assessors, and the authority of the state may be exerted to compel more efficient work on the part of local officials.

Any reform of the property tax requires that assessment be on the basis of the full market value of all taxable wealth and that tax rates be so low as to call for a reasonable, not a confiscatory, contribution from the owners. These two features go hand in hand. Assessed values cannot be brought up to 100 per cent if present high tax rates are to continue, and present rates cannot be materially reduced without increase of the tax base; on the other hand either reform becomes possible if accompanied by the other. The motive to undervaluation through the present method of apportioning state and county taxes must be abandoned; the substitute of the Connecticut method based on the tax revenue is a

simple way out. When once the habit of full value assessment becomes established, review by a board of relief may be made to mean something, and an aggrieved taxpayer may have some real remedy.

These are simple practical changes which any state might introduce, barring certain constitutional limitations which are a handicap to reform in some of the states. Their adoption would mean a tremendous gain in efficiency and equity.

CHAPTER XLVII

THE UNITED STATES REVENUE SYSTEM. THE INCOME TAX. SHIFTING AND INCIDENCE OF INCOME AND PROPERTY TAXES

General property tax and income tax. Just as the general property tax is the mainstay of the revenue system of the American state and local governments, so is the income tax the chief reliance of the federal government of the United States. But this is as far as the historical analogy goes. Whereas the general property tax is one of the very oldest forms of taxation, which long ago reached its prime and which has been on the decline for two or three hundred years in Europe with impending decline even in America, the income tax is a modern tax, developed with astonishing rapidity during the past two or three generations and, in America at least, probably still short of the crest of its wave. Reliance upon the general property tax dates from the first settlement of each of the respective states. The income tax did not become a permanent part of the revenue system of the national government until 1913. As a background for our study of the income tax and the other taxes used by the United States Federal Government, we need to have a general idea of the history of the revenue system of the United States.

Early prejudice against taxation in America. The Anglo-Saxon peoples have always been peculiarly sensitive to taxation. The constitutional history of England, from Magna Charta down to the post-bellum reorganization of today, has been one long series of tax problems. The American colonists were true to the traditions of the mother country, and taxation was one of the sore points between them and Great Britain till it finally merged with other grievances to bring on the Declaration of Independence and the

War of the Revolution. Independence being won, finance was the rock which all but wrecked the new American nation at its launching. The people hated to pay taxes. They had fought to escape the odious taxes of Great Britain. They were now separate sovereign states and, in setting up their national government, they were very loath to give it power to impose upon them the very taxes which they had thrown back in the teeth of Great Britain. It took the lesson of the collapse of the Continental Congress and the government under the Articles of Confederation to force them grudgingly to realize that the taxing power must be granted or there could be no national government at all. It is here that we find explanation of the important restrictions which hedge about the taxing power of our national government as contained in the Constitution of the United States.

Constitutional restriction of direct taxation. The most important restriction is the following (Art. I, sec. 9): "No capitation, or other direct, tax shall be laid, unless in proportion to the census or enumeration hereinbefore directed to be taken." This means that the amount collected by any direct tax must be apportioned among the states in proportion to the population of the several states. For example if the population of state A is double that of state B, the contribution of the residents of state A under any direct tax must be double the amount contributed by the residents of state B. This proportion must hold regardless of the wealth or incomes of the people. It might happen for example that B was a wealthy state and A a poor state, so that while A's population was double B's, the wealth of its residents was half that of the residents of B. Then, to carry out the constitutional rule of apportionment, the rate of the tax on the wealth of the residents of state A would have to be four times the rate imposed in state B. This rule is so flagrantly unjust that Congress has never dared attempt to collect any large amount by means of direct taxes. Some timid experiments were made soon after the adoption of the Constitution and again at the time of the Civil War. But these direct taxes were quite ineffective and produced little revenue, and since the Civil War no attempt to use direct taxes has been made by Con-

gress. The Constitutional rule is practically equivalent to a prohibition of direct taxation.

Indirect taxation: customs and excise. Excluded thus by the Constitution from the field of direct taxation, the national government has had to resort to the less unpopular resource of indirect taxation. From the first revenue act in 1789 down to the Civil War, the customs tariff — *i.e.*, duties on imported merchandise — furnished almost the exclusive source of tax revenue, although from time to time there was considerable non-tax income from the sale of public lands. The only other important form of taxation used was the excise, or duties on the domestic manufacture or sale of merchandise. Alexander Hamilton, foreseeing clearly the weaknesses of the tariff as an exclusive source of national revenue, urged upon Congress the necessity of some sort of internal revenue and succeeded in 1791 and 1794 in securing the adoption of a limited and timid excise upon certain commodities, principally whisky and other distilled spirits. Hamilton's excise taxes were extremely unpopular, actually leading to armed revolt in the "Whisky Rebellion" of 1794 in Pennsylvania. The revenue obtained was insignificant, and the whole system was swept away when the Republican Party succeeded the Federalists in 1801.

During the War of 1812 the breakdown of the tariff system occurred, and necessity forced the country to resort once more to internal excise taxation. The new taxes came too late to be of much service during the war, and they were quickly repealed after its close. That was the end of excise taxation till the Civil War.

The Civil War and the national tax system. The Civil War was a tremendous struggle, and its financial demands were enormous. As had happened in the War of 1812, the tariff system proved a broken reed, quite incapable of serving the necessities of war. Moreover the war was of such duration as to give the people time to overcome their traditional repugnance to internal taxation. So Congress and the people finally learned the lesson which Alexander Hamilton had tried to teach them two thirds of a century before. Starting in 1862, excise taxes were enacted which finally laid burdens upon almost every article of industry and

commerce. The details of the Civil War excise taxes do not concern us here. The significant thing is that, in the financial reorganization after that war, the excise system, though greatly reduced by freeing the majority of taxed commodities, was not entirely abandoned as it had been after the War of 1812. A group of excises, principally upon liquors and tobacco, remained and provided not only an important source of peace-time revenue but also the skeleton upon which it was possible to build quickly an effective system of war taxes in time of future need, of which good use was made on the occasions of the Spanish War and the World War.

Tariff and excise since the Civil War. For the half century following the close of the Civil War, the tariff and the excises have shared about equally (tariff revenue being generally a little greater than excise revenue) the task of producing the revenue needed by the federal government, no other form of taxation, and indeed no other source of revenue of any sort, being of comparable significance. For example during the generation ending in 1913, the combined receipts from customs and excise provided more than 90 per cent of the net ordinary revenue (*i.e.*, excluding loans) of the national government in every year except one.[1]

Beginnings of the income tax. The next revolutionary change in the federal revenue system came with the introduction of the income tax in 1913. Although this marks the beginning of the income tax as a permanent part of the national tax system, it was not our first experience with income taxation. Omitting for the present the discussion of state income taxes, it should be noted that an income tax was first imposed by the Government of the United States at the time of the Civil War. This was a war measure, and its details, though of great historical interest, are not of importance for our present purpose. It was repealed in 1872.

The income tax and the Constitution. In 1894 Congress once

[1] In 1898 this percentage was reduced on account of a large fortuitous income from sale of public property in the Kansas Pacific Railroad and the Union Pacific Railroad.

more enacted an income tax, this time obviously not for war needs. The tax was declared unconstitutional by the United States Supreme Court before it went into effect. Though thus abortive as to practical results, the income tax of 1894 raised constitutional questions which are of the utmost importance to the student of American taxation. As we have learned, Congress is forbidden by the Constitution to levy any direct tax unless apportioned among the states according to their population. Its opponents claimed that the income tax was a direct tax and that, since the law made no provision for its apportionment according to the constitutional rule, it was therefore unconstitutional.

Now the income tax, as regards its economic character, is certainly a direct tax, in that its burden is normally borne by those on whom it is imposed and not shifted to others. But this was not the idea of a direct tax in the minds of the framers of the Constitution, who apparently were thinking only of land taxes and poll taxes. The income tax was accordingly considered an indirect tax, legally, and it was so held by the Supreme Court of the United States when this question was raised with regard to the Civil War income tax. In 1894 however the Court reversed its previous ruling and in the case of *Pollock v. Farmers' Loan and Trust Co.* (157 U. S. 429) held that the income tax was a direct tax, thus bringing the legal definition into practical harmony with the economic concept of direct taxation, though by means of a line of argument which will scarcely stand the test of economic analysis. Being a direct tax, the income tax of 1894 was of course unconstitutional since it was not apportioned among the states according to population. Such an apportionment, as we have seen, would have been intolerable. This decision thus practically deprived the federal government of the power to use the income tax.

The Sixteenth Amendment and the triumph of the income tax. However the popular demand for this form of taxation was so strong that the constitutional obstacle was not permitted much longer to continue. In 1909 Congress imposed a tax upon the incomes of corporations only, this tax by a twist of legal reasoning being regarded as an excise — hence an indirect tax — and so not

requiring distribution according to the population of the states. At the same time Congress proposed an amendment to the Constitution, which was finally ratified by the requisite number of states and promulgated as of March 1, 1913. This amendment, without reopening the question of definition or altering the constitutional rule of apportionment for other direct taxes, expressly gives to Congress the power to impose an income tax without the necessity of apportionment. The ratification of this amendment was immediately followed by the enactment of the income tax of 1913, which imposed a tax upon the incomes of individuals and corporations, absorbing the existing corporate income tax of 1909. This original act has since been revised and extended by various acts, greatly influenced by the World War of 1914–18 and the subsequent financial reorganization. The income tax is now firmly established as a permanent part of the national tax system.

The present revenue system. In the fiscal year 1925 (that is, the twelve months from July 1, 1924 to June 30, 1925) the ordinary receipts of the United States came from the following sources.[1]

Customs	$ 548 million	14.5%
Internal revenue (miscl.)	727 million	19.2%
Income tax	1,761 million	46.6%
Estate Tax	101 million	2.6%
Other ordinary receipts	643 million	17.0%
Total ordinary receipts	$3,780 million	100.0%

Nothing could show more clearly the revolution which the revenue system of the Federal Government has undergone. Up to 1913, as we have seen, the national government had relied for its revenues almost exclusively upon indirect taxation, leaving the whole field of direct taxes to the states and their local subdivisions. Today the Federal Government has not only invaded the direct tax field; it has made direct taxation its chief reliance. Almost half of the total ordinary receipts comes from the income tax. Barely over one third comes from the two great indirect taxes, customs and

[1] *Annual Report of the Secretary of the Treasury on the State of the Finances*, 1925, pp. 141 and 375.

excise, which for a generation before 1913 produced more than 90 per cent of the ordinary receipts. Note also the decline of the customs revenue in comparison with the receipts from excises.

The income tax of 1926. Much historical interest attaches to the details of the income taxes of the Civil War, of 1894, and of 1913, as well as to the modifications made since 1913, particularly in connection with the extraordinary demands placed upon the income tax during the World War and with the reconstruction problems following. The limits of this book however require us to pass at once to a description of the more fundamental features of the present income tax, as contained in the revenue act of 1926.

Individual incomes. The law distinguishes between individuals and corporations. As to the former, all residents of the United States are in general subject to the tax upon all their taxable incomes. Non-residents are generally taxable upon income derived from sources in the United States. The tax is upon the *net income*, which as defined in the law is determined in the following manner. The taxpayer's *gross income* consists of all receipts except those specifically excluded by the law. The principal categories thus excluded are the proceeds of life insurance policies, gifts, bequests, interest on state and municipal obligations, and (to a certain limited extent) interest on United States bonds and notes. Items so excluded need not be reported at all in the taxpayer's return. From the gross income the law permits certain *deductions*, which include the expenses necessary to obtain the gross income, together with taxes, interest, and certain losses and contributions. What remains after subtracting the deductions from the gross income is the *net income*.

The tax rates. The individual tax is in two parts, a *normal tax* and a *surtax*. Before calculating the normal tax, the taxpayer is permitted to subtract from his net income certain *credits*, of which the principal ones are: (1) a "personal exemption" of $1,500 for a single person or $3,500 for a married person or a single person who is head of a family (the combined personal exemption of a husband and wife may not exceed $3,500, which amount may be

taken by either one or divided between them), (2) a credit of $400 for each child under 18 years of age or each person incapable of self-support who is supported by the taxpayer, and (3) dividends received from a corporation subject to the income tax (to be explained below). Upon the base thus determined the normal tax is calculated at the following rates: 1½ per cent on the first $4,000, 3 per cent on the next $4,000, and 5 per cent on all above that.

The base of the surtax is the net income, without deduction of the credits allowed in connection with the normal tax. The surtax rates run from 1 per cent to 20 per cent, as follows:

On the first $10,000 of the net income,	no surtax
On the part of the net income in excess of $10,000 up to and including $14,000,	1%
On the part of the net income in excess of $14,000 up to and including $16,000,	2%

In this fashion the rate increases 1 per cent with each increase of $2,000 in the net income until a net income of $24,000 is reached, upon which the rate is 6 per cent. Thereafter the rate increases 1 per cent with each increase of $4,000 in the net income up to a net income of $64,000, upon which the rate is 16 per cent. The rest of the scale is as follows:

On the part of the net income in excess of $64,000 up to and including $70,000,	17%
On the part of the net income in excess of $70,000 up to and including $80,000,	18%
On the part of the net income in excess of $80,000 up to and including $100,000,	19%
On the part of the net income in excess of $100,000,	20%

Earned incomes. In the tax on individual incomes the law makes a certain rough discrimination in favor of earned incomes as against those that proceed from property or investments. The provisions for accomplishing this end are as follows: if the taxpayer's net income is less than $5,000, the entire net income is assumed to be earned, regardless of the source in the particular case, and if the net income is more than $5,000, at least $5,000 is assumed

to be earned. Earned net income in excess of $5,000 must be proved, and the total earned income of any taxpayer is limited to $20,000, regardless of the source. These provisions are obviously quite arbitrary and can result in only a very rough approximation to the actual earned income in any given case. They were so drawn for the sake of administrative simplicity. The taxpayer's earned net income being thus determined, he is entitled to a reduction of the amount of his tax by 25 per cent of what his tax would have been if his earned net income had been his only net income, provided that this deduction shall not be more than 25 per cent of the normal tax on his actual income plus 25 per cent of the surtax on the earned income.

Income of corporations and partnerships. Corporations are taxed 13½ per cent[1] upon their net income and are not subject to surtaxes. The dividends received by an individual from a corporation subject to the income tax are deducted from the individual's net income in computing his normal tax, but are not deducted in computing the surtax. This accomplishes a partial correction of the double taxation which results from taxing individuals upon dividend income which has already been included in the base of the corporation income tax.

Partnerships as such are not taxable on their income. Their net income, whether actually distributed or not, must be allocated to the several partners and included by each in his individual income.

An example. The important features of the income tax and the technical terms employed may be made clearer by a practical illustration. Let us take the case of a lawyer whose total receipts during the year 1925 consisted of: professional fees, $20,000, dividends from New York Central Railroad stock, $4,000, and a bequest of $10,000. He paid out during the year for professional expenses, $5,000, taxes, $1,000, and contributions, $1,000. He was married and had 4 children aged respectively 7, 10, 15, and 19 years. His income tax for the year 1925 would be determined by the following computation:

[1] This is the rate in effect after 1926. The tax paid in 1926 on incomes earned in 1925 is at the rate of 13 per cent.

Professional fees,			$20,000	
Dividends from N. Y. C. R. R.,			4,000	
Total *gross income*,				$24,000
Professional expenses,			5,000	
Taxes,			1,000	
Contributions,			1,000	
Total *deductions*,				7,000
Net income (base for surtax),				17,000
Personal exemption,			3,500	
Credit for 3 children,			1,200	
Dividends,			4,000	
Total *credits*,				8,700
Base for normal tax,				8,300
Normal tax:	on $ 4,000	@ 1½%	60	
	on 4,000	@ 3%	120	
	on 300	@ 5%	15	
Total *normal tax*	on 8,300			195
Surtax:	on 10,000		no surtax	
	on 4,000	@ 1%	40	
	on 2,000	@ 2%	40	
	on 1,000	@ 3%	30	
Total *surtax*	on 17,000			110
Total normal tax and surtax				305

The earned net income in this case is evidently the $20,000 of fees less the professional expenses of $5,000, or $15,000, which is clearly within the $20,000 limit of the law. The tax upon this earned net income (determined by a calculation similar to the above) would be: normal tax, $295, plus surtax, $60; total, $355. Twenty-five per cent of this is $88.75, but the deduction actually allowed is limited to an amount computed thus:

Normal tax on actual income, $195, of which 25% is	$48.75
Surtax on earned income, $60 " " " "	15.00
Allowed deduction on account of earned income	$63.75

The final step in the computation of the tax is as follows:

Total normal tax and surtax	$305.00
Deduction on account of earned income	63.75
Tax due	$241.25

The reader will of course understand that the foregoing is only the briefest sketch of the most essential features of the United

States income tax. The law itself is a lengthy document equivalent to about 100 ordinary printed pages, loaded with technical details and administrative provisions of the utmost complexity. To master all the details of this law, with the mass of administrative rulings and court decisions which has grown up about it, is an ambitious task to be undertaken only by the special student of taxation. We have now before us however sufficient knowledge of the law to enable us to proceed to the discussion of the more fundamental principles involved, in connection with which we may draw some comparisons with the income tax systems of other countries.

The exempt minimum. The minimum amount of income which should be exempt from taxation is governed by two important principles. There is on the one hand the generally admitted advantage of having the income tax as widely distributed as possible; this would indicate the smallest possible exempt minimum, limited only by the bare "minimum of subsistence." On the basis of this principle alone the personal exemptions and credits for children and dependents in the United States law are rather high; they are considerably higher than the exemptions granted by the tax laws of most other countries. In the original law of 1913 the exemptions were set at $3,000 for a single person and $4,000 for a married person. These figures were extraordinarily high; they certainly did not even approach the minimum of subsistence; and they relieved from the income tax the great majority of the people. The total number of personal returns for 1916, after the law had been in operation three years, was only 437,036. This was less than half of one per cent of the population of 101,000,000, and a considerable number of these had no taxable income and paid no tax. The tax was actually a tax on only a few wealthy or well-to-do persons.

The exigencies of the World War compelled the reduction of the personal exemptions to $1,000 and $2,000 respectively, though the latter limit was increased to $2,500 in the revisions of 1921 and 1924. The number of individual returns increased eight-fold in the next year, being 3,472,890 for the year 1917. For the year 1923 the number of returns was 7,698,321, nearly eighteen times

the number reporting for 1916. Returns for 1923 were filed by 7 per cent of the population, as compared with less than $\frac{1}{2}$ of 1 per cent for the year 1916. In the revision of 1926 the limits were further increased to $1,500 and $3,500 respectively. The reader will of course recognize that the figures above do not mean that the income tax reached only 7 per cent of the people. This is the percentage of returns to the whole population. But the majority of the population would not make returns even if there were no exempt minimum. For most married couples only one return is made, and no returns are ordinarily made by minors and other dependents. Nevertheless, the income tax still reaches only the minority of the people, and this will probably always be the case, for reasons which will presently be made clear.

The principle of wide distribution is opposed by a contrary principle. The attempt to bring the income tax down to the smallest incomes which can possibly bear it is accompanied by expenses which will ultimately approach the revenue to be obtained from the smallest incomes. To collect a few dollars, or only a few cents, from each one of thousands of taxpayers is not profitable nor justifiable if the revenue thus collected is not materially in excess of the cost of collection. In general the limit thus indicated will be reached long before the tax encroaches on the bare minimum of subsistence, and we are therefore justified in reconciling these two principles by stating that the purpose of the income tax will best be served when the exempt minimum is made as small as is consistent with administrative efficiency and fiscal profit. In the light of this conclusion, the existing personal exemptions and credits for dependents appear disproportionately high.

What is taxable net income? Since income is the measure *par excellence* of tax-paying ability, the broadest possible definition of income is logical. This rule is followed on the whole in the United States tax. The definition of gross income is broad and inclusive, and the exclusion of such items as bequests, gifts, and the proceeds of life insurance policies is logical. The obvious deductions from gross income are those expenditures which are necessarily incurred in obtaining the gross income. The deduction of taxes and interest

involves no serious question of principle. Deduction of specified contributions, though more questionable, is usual. The two most serious problems of taxable income in connection with the United States income tax are (1) the exclusion from gross income of interest on State and municipal and certain federal obligations and the salaries of state and municipal employees and (2) the treatment of income from investment in the stock of domestic corporations.

Tax-free securities. The law specifically exempts from taxation interest upon bonds of states and local governments, interest upon bonds of Federal Farm Loan Associations and Federal Land Banks, dividends of Federal Land Banks and Federal Reserve Banks, and, to a limited extent varying with the several issues, interest on bonds of the United States. This sets up a large and rapidly growing class of income which may be received without incurring tax obligation and invites investment in such obligations in order to escape taxation. Because of the extraordinarily high and steeply progressive rates of the income tax, the seductiveness of this invitation increases in direct proportion to the wealth of the taxpayer. The investor whose net income is under $10,000 has no surtax to pay, and the highest taxed increment of his income pays only the normal tax of 5 per cent. The advantage of a tax exempt security is of course something to him but it is not great. He would not, for example, choose a 5 per cent municipal bond in preference to a $5\frac{1}{2}$ per cent industrial bond otherwise equally desirable. To the very wealthy the inducement is far greater. A taxpayer with a net income in excess of $100,000 pays 25 per cent tax on the part of his taxable income in excess of that figure. Investment of that part of his income in a 5 per cent tax-free bond will be more profitable than any other investment, equally good, which pays less than $6\frac{2}{3}$ per cent. Below this limit the margin of preference of tax-free over taxable securities varies according to the taxpayer's income and the amount of his investments. During the World War and for half a dozen years thereafter the rates of the individual income tax were extraordinarily high, the maximum surtax rate being 65 per cent in the law of 1918, 50 per cent in 1921, and 40 per cent

in 1924. These high rates further enhanced the attentiveness of the tax-free securities to the wealthy taxpayers.

The volume of tax-free securities. Coincident with the preference thus created, there was a great increase in the volume of tax-free securities offered to the wealthy investors. As we have seen, the United States debt has increased from about a billion and a quarter to over twenty billions, though this debt enjoys only a limited exemption from income tax. The total volume of state and municipal bonds has grown with extraordinary rapidity. Before 1919 the total annual issues had never equalled half a billion dollars; in each of the three years 1921 to 1923 the issues have exceeded a billion. The United States Treasury Department estimated that the amount outstanding on January 1, 1924, was $12,309,000,000. There have also been the large issues of bonds of the Federal Farm Loan Associations and the Federal Land Banks.

Escape of the wealthy taxpayers. The evidence at hand is not sufficient to tell us exactly to what extent this door of escape was availed of by wealthy taxpayers. Significant evidence comes from the probate court records of the settlement of large estates. Also noteworthy is the decline in the relative number of very large incomes reported for taxation. The statistics of income tax returns show that the number reporting net incomes of all sizes up to $30,000 increased 123 per cent from 1917 to 1923, whereas the number above $30,000 increased only 7 per cent. Those reporting incomes above $100,000 declined by 37 per cent, and the number above $500,000 was reduced from 456 to 215, a decrease of more than half. This decline of the large taxable incomes is due to a variety of causes: probably the most important is the state of business depression which prevailed after 1920. Also wealthy taxpayers were exercising more care to make their tax obligation as small as possible in lawful ways, and it is probable that actual evasion had increased, particularly since the patriotic restraint of the war period was removed. There is no doubt however that one cause was the escape of wealthy taxpayers through the inviting door of tax-free securities. This state of

affairs was contrary to the spirit of equitable taxation and threatened to discredit the whole income tax system, until the situation was in part relieved by the lowering of the surtax rates in 1926.

What the government gains from tax-exemption. It may be asked, why does the law grant exemption to interest on government securities with results so obvious and so unfortunate? There are two principal reasons. In the first place tax exemption, by giving an added attraction to the favored securities tends to make it possible for the government to borrow at somewhat lower rates of interest. In the simplest case, assuming all disturbing circumstances eliminated, the gain to the government from borrowing at the lower rate would just equal the loss from tax exemption. Thus suppose the government issues $1,000,000,000 of bonds which may or may not be made subject to a permanent income tax at the flat rate of 10 per cent. If the government could float such an issue at par, bearing a rate of interest of 5 per cent and subject to the income tax, it is obvious that the bonds could be floated at $4\frac{1}{2}$ per cent if exempt from income tax, provided that all the bonds were destined to go to persons subject to the income tax. It is said then: why go to the trouble of taking in taxation what must be given back in a higher rate of interest? This reasoning has been generally accepted and has led to the common practice of governments exempting their own securities from both income and property taxes (for the principle is the same for both kinds of taxation).

The effect of progression. It is a mistake however to assume that any such simple relation between tax exemption and interest rates exists under the complex circumstances of the United States income tax. The chief disturbing factor is the presence of progressive rates. The privilege of tax exemption will be variously valued according to the investor's total income otherwise taxable, and the government's gain in interest rates will be determined by the value placed on tax exemption by those in the lowest income class which must be appealed to in order to market the entire loan. Such large loans as were made by the United States during the

World War could not be placed entirely with the very wealthy, and in fact urgent appeal was made to the whole public, including those of very moderate means. To such the tax exempt feature was of slight value. The government cannot have different interest rates for different classes of investors; hence the possible gain to it is limited by the slight advantage that accrues to investors of moderate means. By the same token the gain from tax exemption to the wealthy is very material, only slightly impaired by the small possible reduction in the rate of interest on government bonds. Stated in general terms we may conclude that, when a government having a progressive income tax makes a large loan, the rate of interest which it has to pay will be only slightly reduced by making the income from the bonds tax-free, and its loss from tax exemption will be greater than its gain from a lower rate of interest. This conclusion has been recognized by a number of governments. The British loans for example have generally been made without tax exempt features.

State and municipal bonds and the Constitution. If under modern circumstances the advantage to a government from exempting its own bonds is small, there is of course no gain at all from exempting the bonds of another government. The United States government had nothing to gain from exempting the interest on state and municipal obligations. Here the reason was a legal one. It is commonly believed that a tax imposed by the United States government upon the income from interest on obligations of the states or their minor subdivisions would be unconstitutional. There is good reason to question the accuracy of this opinion, but Congress did not wish to risk the overthrow of the war taxes and so avoided the issue by granting the exemption. Since the war the serious results of this policy have been apparent, but Congress has been unwilling either to pass a law removing the exemption or to recommend a constitutional amendment settling the question. Of course exemption once granted for loans already made can not be withdrawn without breach of good faith; here the mischief has been done and cannot be undone. For future issues the remedy is possible, and the longer it is delayed the more serious the condition

becomes with each addition to the growing volume of state and municipal bonds and other tax-free securities.

Salaries of public employees. Salaries of employees of the United States are generally subject to the federal income tax like any other salaries, which is logical, but the salaries of state and municipal employees are exempt. The reason for this otherwise illogical exemption is the same constitutional restriction which is believed to prevent the taxation of interest from state and municipal securities; as to the correctness of this interpretation of the Constitution there is the same question.

Dividends of domestic corporations. The treatment of dividends from domestic corporations is peculiar. As we have seen, corporations are taxed upon their net income at the rate of $13\frac{1}{2}$ per cent, and in consideration of this the stockholders are relieved of the normal tax upon dividends received, though they must pay the surtax. This arrangement is on the face of it a discrimination against income from corporate dividends, since the tax on corporations is considerably heavier than the normal tax which is thereby remitted; and a closer analysis will show that the discrimination is even more severe than may at first appear.

Let us take the case of a married person with two minor children having no income except his dividends upon stocks in a domestic corporation. Suppose that his share of the net income of the corporation available for dividends before paying the income tax were $5,000. The corporation pays income tax at $13\frac{1}{2}$ per cent upon its net income, which of course correspondingly reduces the amount which it can distribute in dividends to its stockholders. Our individual's burden on account of this tax is $13\frac{1}{2}$ per cent of $5,000, or $675. Now supposing his income had come from other sources, what would his tax have been upon a net income of $5,000? He is entitled to the personal exemption of $3,500 and to credits of $800 for his two children. This leaves only $700 taxable, at the normal tax rate of $1\frac{1}{2}$ per cent, and his tax is $10.50, instead of $675. If his income had been $4,300 or less, he would have paid no tax at all, whereas a portion of corporate income amounting to $4,300 would be reduced by a tax of $580.50 before being paid out

in dividends. Our method of taxing corporate earnings, through the corporation instead of the individual stockholder, thus deprives the stockholder who does not have considerable other income of the advantages of the personal exemption and credits and of the lower rates of normal tax to which the spirit of the law entitles him, and the magnitude of the injustice varies inversely with the taxpayer's income. The remedy would be to exempt corporations from tax on that part of their net income distributed to stockholders who were themselves subject to the income tax and then to treat individual income from dividends exactly like other kinds of income.

The thing which chiefly distinguishes the income tax is its personal nature and its close relation to ability to pay. These characteristics are to be obtained in the fullest measure only when the definition of taxable income is as broad as possible. Whatever narrows the concept of personal taxable income, whether by exemption or stoppage at source or other devices, tends to weaken the income tax.

Income tax rates. The most significant feature of any tax is of course the rate. A survey of income tax history discloses great variety of rates. The oldest income tax in the world is that of Great Britain, which was first imposed in 1798 and has been a permanent part of the British tax system since 1842. The basic rate of the British income tax has generally been moderate; for example in the fifteen years before the World War the rate (stated as so many pence to the pound) varied between $3\frac{1}{3}$ and $6\frac{1}{4}$ per cent, the highest rate being in effect for only one year (1902–1903) during the South African war. A tendency to gradual increase of the rate is evident. Up to 1910 the British tax was imposed at a flat rate without progression; from 1910 to 1914 the super tax was only 6 pence or $2\frac{1}{2}$ per cent added to the basic rate. In Germany the income tax, introduced by most of the states of the Empire during the last part of the nineteenth century, was at progressive rates generally ranging between $\frac{1}{2}$ of 1 per cent and 5 per cent. In the United States the Civil War income tax was imposed at roughly progressive rates with a maximum of 10 per cent. The rate of the

abortive income tax of 1894 was 2 per cent and applied only to incomes in excess of $3,000 or $4,000. The present income tax was inaugurated in 1913, with progressive rates from 1 to 7 per cent.

High rates due to war. The epoch of moderate rates was brought to a sudden close by the World War. The scale of surtax rates of the United States income tax in 1919 was from 1 to 65 per cent, which, added to the normal tax rates of 6 or 12 per cent, made a total rate of 77 per cent upon that part of an individual's income in excess of $1,000,000. This is the highest tax rate known to modern history. The British tax was increased only slightly less; the highest rate was lower than in the United States, but the rates upon the middle incomes were higher than ours.

Such rates as these would of course never have been seriously considered except for the extraordinary demands of the war, and even so they were probably pushed beyond the limits of maximum return. However war excitement and patriotic fervor made their enactment possible and their collection not entirely impracticable. Since the war the impracticability of such rates has been fully demonstrated in the United States. Even with the best possible administration, which we in the United States confessedly do not have, the success of an income tax must depend very materially upon a prevailing spirit of fairness and coöperation on the part of the taxpayers. So long as this spirit is general, the law and its administrators can cope with the actual criminal evaders. When this spirit disappears, when there is a general feeling that only the quixotic few pay their full taxes, when the majority become law-breakers, then no administrative machinery that human intelligence can devise will be able to master the problem. This condition is threatened if rates are made too high.

There is another aspect of high rates which is deserving of mention. The income tax is peculiarly an elastic source of revenue; its yield may be varied by raising or lowering the rates. As such it has been a tower of strength in time of war or other national emergency. But unless the rates are reduced to moderate limits after the emergency, there will be no elasticity left for the next emergency.

Stoppage at source. As regards the administration of the income tax, two rival principles have been evolved. The first, which may be called the principle of *stoppage at source*, has been exemplified in the British income tax. All incomes are classified into five schedules for the purpose of making the utmost use of this principle. For example a landlord is not called upon to pay tax upon his rent; that tax is collected from the tenant, who thereafter deducts the tax from the amount of his payment of rent to the landlord. In like manner the income tax upon interest, dividends, salaries, and various other forms of income is already paid and deducted before the income comes to the recipient. Only where such stoppage is not possible, does the British law require a declaration and payment by the taxpayer himself.[1] This principle has two great advantages: it minimizes the motive for evasion, and it preserves to the greatest possible degree privacy as to the amount and character of each individual's income.

Personal declaration. The other principle of income tax administration may be called that of *personal declaration*. Germany has been the principal exponent of this idea. The taxpayer is required to make a return disclosing in considerable detail his taxable income from all sources, and it is the business of the administrative officials to check up and verify these returns and then to collect from the taxpayer the amount due. Stoppage at the source, if availed of at all, is little used. This principle foregoes the very important advantages of the British method, but it has the one great advantage of making progressive rates possible. Stoppage at the source without personal declaration is inconsistent with the principle of progression. Thus Great Britain, having the oldest and best organized income tax in the world, clung to her system of a uniform flat rate long after most other countries had been won over to the idea of progression. It was not till 1910 that she permitted the first breach in her traditional policy by adding a super tax in the case of incomes in excess of £5,000, although in 1907 the introduc-

[1] This description relates to the British income tax as it long existed before certain recent changes which will be noted in the following paragraph.

tion of differentiation in favor of earned incomes had involved the personal declaration of those claiming the advantage of differentiation. Finally the World War compelled Great Britain to embrace completely the idea of progression, with general requirement of personal declaration and the sacrifice of many of the characteristic features of her former system.

American practice. In the United States the system of personal declaration has always been followed, with varying and rather haphazard use of stoppage at source. The limited possibilities of stoppage at source in connection with a progressive tax, together with the weakening of the personal nature of the income tax which is apt to result (cf. our discussion above of the treatment of income from dividends), have gone far to diminish the favor formerly accorded to the principle of stoppage at source. It has been learned that most of the administrative advantages of stoppage at source could be obtained by requiring reports from those making payments of rents, dividends, interest, salaries, etc., without requiring the deduction of the tax itself. There has thus developed in America the new principle of "information at source," which appears fairly well to embrace the advantages and avoid the disadvantages of stoppage at source in connection with a progressive income tax.

State income taxes. The income tax is not necessarily a national tax only. In Germany, before the revolution of 1918, the income tax belonged exclusively to the separate states composing the imperial federation. In the United States the income tax was used by a considerable number of the states before the Federal Government took its first step in this field. Some state income tax laws were enacted between 1840 and 1850, and during the second half of the nineteenth century the number grew to sixteen. These laws all carried over enough of the administrative features of the general property tax to insure their complete failure. Assessment was based upon the personal declaration of the taxpayer, and administration was in the hands of local officers; by the end of the nineteenth century the breakdown was even more complete than has been exhibited by the general property tax.

A new epoch in state income taxation began in 1911 with the enactment of the Wisconsin tax, a carefully drawn law providing for complete administration by state officials. It has been a distinct success. Since then a dozen other states have enacted personal income tax laws, which, profiting by the experience of the earlier epoch, generally provide for such centralized state administration as to give them at least a fair chance of success. Among the more conspicuous examples are those of New York and Massachusetts. A few states also levy income taxes upon corporations only. Having already investigated the more important principles of income taxation, we shall have to forego further study of these state taxes. It is likely that the future will see a still further development of successful state income taxes, a movement which receives impetus from the impending disintegration of the general property tax.

Shifting and incidence of taxes on property and income. Wealth gives rise to income, and the value of wealth is the discounted value of its expected future income. This relation between wealth and income, developed in an earlier part of this book, makes the property tax and the income tax very similar as regards their shifting and incidence and is our reason for postponing the study of the shifting of the property tax to this point, where it may be studied in connection with the shifting of the income tax. We have already learned, in a general way, that shifting of the burden of a tax by the immediate payer to someone else is conditioned upon a power, given by the tax itself, to raise the price of some commodity or service which the immediate payer has to sell. Since the general property tax and the income tax reach various kinds of property and income, the study of their shifting and incidence can most conveniently be broken up into more or less separate problems dealing with the more important classes of property and income. We need then to inquire in each such case to what extent the power to increase a price is given by the tax and what results follow when such a power is given.

The tax on economic rent. Let us consider first taxes upon land and the income from land. The simplest theoretical case is that

of the tax upon the landowner's income from economic rent. The reader may find it convenient at this point to refer back to Chapters XXXIII and XXXIV, which investigated the nature of economic rent. Having in mind the conclusions there reached, let us ask what would be the result of a 10 per cent income tax upon the economic rent of $1,000 received from 50 acres of agricultural land. The landowner's income from this source is reduced by $100; will he recoup himself by passing the tax on to someone else? If the land is rented to a tenant, his only possible recourse is to increase the rent. But assuming, as we have a right to do, that the tenant is already paying the full economic rent, he can pay no more without being forced to operate at a loss. Rather than do this, the tenant will leave this particular land and either go out of business or betake himself to the marginal land, where he can do as well without the payment of rent and of course without the burden of any tax on rent. Since rent is the result of superiority over the marginal or no-rent land and since the profitableness of operations on the marginal land is in no wise affected by a tax on economic rent, the tax can have no effect upon the amount of economic rent that can be exacted for the use of the superior land.

But, it may be objected (by someone who has not perfectly recalled the theory of economic rent), cannot the tenant, when called upon to pay higher rent, in his turn make good his loss by raising the price of his product? The answer is no. The price of agricultural products is governed by demand and supply. Certainly there is no conceivable way by which a tax on economic rent could affect the demand for agricultural products. On the other hand the supply is governed by the cost of production on the marginal land, which pays no rent and so pays no tax. In other words, economic rent does not enter into the price of the products of land; and a tax upon economic rent, since it cannot affect either the demand or the supply of such products, gives no power to raise prices. By the same token, if the landowner is himself using his land, he has no more power than a tenant to shift a tax on his economic rent by raising the prices of his product. We conclude then that a tax on the economic rent of land, giving the owner no

power to raise any price, cannot be shifted but will be borne by the immediate payer, the landowner. And this is true, no matter how heavy the tax, up to the limit of 100 per cent, taking the entire economic rent. To this extent the theoretical basis of the " single tax " is unassailable.

The tax on land values. We necessarily reach the same conclusion as to a property tax upon land alone, since the value of the land is the capitalized value of its economic rent. A property tax on land can certainly give the owner no power to sell the land at a higher price; nor does it give him power either to exact more rent from a tenant or charge higher prices for agricultural products, since the marginal land, having no economic rent and so no value, is no more affected by a property tax than by a tax on rent.

The practical result. The foregoing analysis is of course highly theoretical, but it is not for that reason inapplicable to the practical problems of taxation. For the various disturbing factors which may interfere with the practical working of the theory of economic rent, the reader is referred to the previous chapters on that subject. For our present problem, it may be observed that the disturbing factors are mainly two. In the first place it is quite possible that landlords may be easygoing or careless and may not be exacting the full economic rent from tenants. A tax may open their eyes, overcome their tenderheartedness, and lead them to exert every effort to obtain all the economic rent. The result of the tax may thus be an additional burden upon the tenant. Note that this is no exception to the principle that the tax gives no power to raise rents. In this case the landlords have simply been aroused to exert the power which they already had to exact the entire economic rent.

The other disturbing factor comes from the fact that taxes are actually imposed upon property and income, without distinguishing either pure land value or pure economic rent from other forms of real estate or income. Rents actually received may include income from structures on the land, taxes upon which may possibly be subject to shifting, and real estate is frequently valued for tax purposes as a whole, including the value of the structures with the pure land value. These facts are no denial of the truth of the

theory; they simply serve in some cases to obscure the actual results. In general the practical results are as indicated by the theory; taxes on the income from land and on land values are not shifted either through higher rents or through higher prices of the products of land.

Capitalization of land taxes. By what is known as the process of "capitalization," taxes on economic rent or land values tend to rest as a permanent burden upon those who owned the land at the time of the imposition of the taxes and not to be passed on to later purchasers. If an acre of land is yielding an economic rent of $10 a year and the rate of interest at which such incomes are discounted is 5 per cent, the land will be worth $200. If now, assuming no other material change in the tax system, an annual tax is imposed on economic rent at the rate of 10 per cent, the landowner will have left an income of only $9 a year, which, capitalized at 5 per cent, makes his acre of land worth $180. Taking 10 per cent of the annual income is equivalent to taking in one lump sum 10 per cent of the value of the land. No one will buy this land for more than $180, and the purchaser, receiving $10 in rent and paying $1 in tax, obtains 5 per cent on the investment, the same as he would have received if he had paid $200 for an acre not subject to the tax on its rent. Of course precisely the same result follows the imposition of an annual property tax on the value of the land. It is in this sense correct to say that a non-shiftable tax tends to become a "burdenless" tax. What occurs is that the government takes to itself once for all a certain part of the value of the taxed land. Later purchasers pay for only the part that is left, so making allowance for the taxes they will have to pay. Of course any new tax or any increase in the tax rate will burden the then owners of the land.

Taxes on dwelling houses or on their rent. Let us turn our attention now to the effect of a property tax upon dwelling houses, taking first the case of houses which are rented. The rent paid for a dwelling house, as distinguished from the rent paid for the location, is a competitive price determined by the demand and supply of such houses for rent. Now of course a tax on such houses will not affect the demand of tenants. But it will affect the supply.

Before the imposition of the tax, there was a certain normal return obtainable from renting dwelling houses; now that return has been reduced by the amount of the tax. Investors will tend to stop building new houses, preferring to put their capital into other ventures where the normal return is obtainable. The diminished supply of houses for rent, as old houses wear out and are not replaced by new ones, will cause an increase in the rents which house owners can exact and tenants will pay, until rents are high enough to encourage new building. In a stationary population the diminished supply together with unchanged demand will bring about this result as soon as old houses wear out. If the population is increasing, thus requiring new building at a faster rate than the decay of old houses, the increased demand coupled with the decreased supply will bring the result the more quickly. Only when population is declining is it possible that a diminishing demand may offset the decreasing supply due to decay of old houses and so postpone for a time the increase of house rents. Even in this case the increase must come with new building sooner or later, unless the community is destined to complete decay.

All of this will be recognized as merely an application of the simple laws of price with which the reader is perfectly familiar. The same reasoning will show that if the tax were imposed immediately upon the tenants they would have to bear it, having no means of compelling landlords to reduce rents on account of the tax. And finally, so far as houses are occupied by the owners themselves, there is no question of a price to be charged and hence no possibility of shifting. The same results will clearly follow if, instead of a property tax upon dwelling houses, we have a special income tax on house rents. Our general conclusion then is that the final incidence of a tax either on dwelling houses or on their rent tends to be upon the occupiers, whether they be owners or tenants, the tax being shifted when imposed upon the owners of rented houses. What distinguishes buildings from land and causes opposite results as regards the incidence of taxation is that buildings are constantly wearing out and being replaced with new ones, whereas land is permanent and practically fixed as to quantity.

The tax on business property. The property tax on wealth used in business, such as factory and store buildings, warehouses, railway equipment, machinery, goods and materials, etc., generally means an addition to the cost of the products of business enterprise. Since, as we have learned, cost of production is the fundamental factor in the supply of goods and services and has therefore an important part in the determination of prices, anything that increases cost of production will tend to raise prices. If prices should not rise in response to such a tax, some of the producers would eventually be forced out of business and those that remained could exact higher prices. It is thus that a property tax upon wealth, other than land, used in production is generally shifted through higher prices of products and so finally rests as a burden upon the consumers of the products.

We have assumed competitive conditions, and the result will be different in the case of a tax upon an instrument of production used by a monopoly. Monopoly price does not tend to equal cost of production, and a tax which increases the cost of production will not necessarily drive the monopoly out of business or lead to curtailment of its product. It will however change the conditions upon which the monopolist has calculated the price that gives the maximum profit and may therefore lead to the determination of a higher price. In that case, the amount of the product sold will fall off and the monopoly will suffer a loss of profits, though less of a loss than if it had not been able to shift a part of the tax through a higher price. In the ordinary case the tax on an instrument used by a monopoly is thus partly borne by the monopolist and partly shifted to the consumers of the monopoly product.

The property tax on consumers' goods. When the property tax falls upon goods in the hands of the final consumers, such as household furniture, books, jewelry, automobiles, etc., there is no longer any question of a price to be charged and hence no possibility of shifting the tax.

The income tax. The personal income tax is on the whole not much subject to shifting. In the case of the tax on competitive business profits there may be a certain degree of shifting through

the charging of higher prices. This can happen only if the added burden of the tax at least threatens to drive out some of the marginal producers. But since the tax only encroaches upon net income or profit, and takes only a fractional part of that, few producers, and those only the very weakest, are likely to give up the fight while still retaining nine tenths or four fifths or some other substantial fraction of the net income that formerly sufficed to keep them going. The case is quite different from the property tax, on machinery for example, which adds to the cost of production of each unit and may readily wipe out profits and substitute losses for the weaker competitors.

On the other hand an income tax upon the profits of a monopoly cannot possibly be shifted. The monopoly, having already set its price so as to reap the maximum profit, has nothing to gain by a change in price, even though it face a tax up to 100 per cent of its profits. It is better to retain any given fraction of a larger net income than the same part of a smaller income which would result from a change in price.

Upon non-business incomes the possibility of shifting is even less than in the case of business incomes. Ordinarily there is here no power to increase any price on account of the tax and so no possible way of shifting its burden. We are justified in concluding that in general the income tax either is not shifted at all or is shifted only occasionally and in slight degree through an increase in the prices of goods and services.

CHAPTER XLVIII

THE CUSTOMS TARIFF. PROTECTION

The customs tariff. In the brief historical survey of the United States revenue system with which the preceding chapter opened, reference was made to the customs tariff as the most important source of national revenue down to the introduction of the income tax in 1913. From the beginning of the American nation in 1789 to the Civil War, the tariff was almost the only tax imposed by the national government, and from then to 1913 it divided honors about equally with the excise or internal revenue taxes. Today the tariff ranks third among the sources of federal revenue.

Customs or tariff duties are taxes imposed upon those who bring goods across the frontier, the base being the goods so moved. Since goods may be moved across the frontier in either direction, we have import and export duties. The Constitution of the United States prohibits all duties on exports and also practically forbids the levy of any customs duties by the states. In America therefore we are concerned solely with import duties levied by the national government. Other countries do not have such restrictions upon export tariffs, and even in America before the adoption of the United States Constitution the several colonies or states had their own separate tariff systems, in which the other colonies or states were on the same footing as the outside world, and in which export duties were a common incident. As a general rule however duties on exports are sparingly used as compared with import duties, and the most important and interesting problems of the tariff, not only in America but everywhere else, have to do with the tariff on imports.

Technical and administrative features. Unlike the property and income taxes, the general technical and administrative features of the tariff are fairly simple and obvious, and only a few words

need be said on this part of the subject. Tariff duties or rates are either specific or *ad valorem*, terms which we have previously defined. For example the (specific) rate on butter might be 2½ cents per pound, while the (*ad valorem*) rate on woollen cloth might be 35 per cent of its value. Administration of the tariff requires establishment of customs houses at the principal ports of entry and the principal railway entrances into the country, the legal requirement that goods shall be brought into the country only through the customs houses, and enforcement by means of constant patrol of the frontiers to prevent smuggling. Goods when entered in the customs house must be examined and classified as between dutiable and free; dutiable goods must be measured and (in case the duty is *ad valorem*) appraised; the proper legal rate for each article is determined and the amount of the duty calculated and collected from the importer. It is a common practice to provide government warehouses in which dutiable goods may be stored, to be taken out at the convenience of the importer by payment of the duty and other charges.

In some countries, as in Great Britain, the tariff is a simple system of rates upon a few dutiable articles. In others, of which the United States is an example, almost everything is dutiable and the tariff is a bewildering complex of hundreds of rates, specific, *ad valorem*, and in all sorts of obscure combinations. The most difficult part of the customs administration has to do with the appraisal of goods, where there is a problem somewhat similar to that which confronts the assessor of the general property tax. However the problem is here made easier by the concentration of most imports in the large customs houses, where a corps of expert appraisers is constantly employed, and there is no reason why the customs administration should not be fairly effective.

Shifting and incidence. The most interesting problems of the customs tariff have to do with its purpose, its shifting and incidence, and its ulterior economic effects.

The economic principles governing the shifting and incidence of duties on imports are simple, when once the laws of price have been mastered. Let us suppose that a certain kind of pocket knife is

being imported (without duty) and sold in the home market at a price of one dollar. This price, according to the laws of price, tends to be about equal to the cost of production, including of course transportation and all other costs necessary to deliver the article in the market. Now let us suppose that a tariff is imposed at the rate of 25 per cent. Upon each knife entered the importer must pay a tax of 25 cents. This tax will be an addition to the cost of delivering the knife, and the selling price must rise by the amount of the duty to $1.25. The tax is thus shifted by the immediate payer, the importer, to the consumer. Of course if the importer is bringing in an article for his own use, as when American travelers return from Europe with their personal purchases, the duty is immediately on the consumer and no shifting is required. The general rule is that the incidence of a tariff is upon the consumers of the imported goods, the tax being shifted when the importer is not himself the consumer. The possible exception of an article sent into the country by a foreign monopoly is not of practical importance.

Economic effects of the tariff. A tariff duty may be so high as to stop the import of the taxed article, either (1) because the demand is such that none will be bought at the new price, or (2) because it becomes cheaper to produce the article in the home country. If nobody were willing to pay $1.75 for the particular knife used in our previous illustration, a duty of 75 per cent would stop all further import. Also if the cost of making such a knife in the home country were $1.50, any duty in excess of 50 per cent would stop imports, since home manufacturers would sell the knife for less than the new price of the imported article. This tendency to check import of the taxed article and to encourage the establishment of the domestic industry is one of the most important characteristics of the tariff and raises some fundamental problems, among them the time-honored controversy over "protection" to which we shall give our attention later in this chapter.

A numerical illustration. It may help the reader to grasp the characteristics of the tariff which have been pointed out, together with their immediate consequences, if we set up a simple numerical

example. Let us suppose that a certain kind of cloth is being imported, free of duty, and sold in the home market at $1.00 a yard and that 1,000,000 yards are sold in a given year. We assume that the cost of making this cloth at home is $1.50 a yard. Obviously the reason for the import is that this is the cheapest way to obtain the cloth, and of course there will be no home manufacture. The simple situation resulting is shown in column 1 of the table below:

	1 *No Tariff*	*2* *10% Tariff*	*3* *50% Tariff*	*4* *75% Tariff*
Number of yards imported	1,000,000	900,000	300,000	. . . .
Number of yards domestic	. . . . .	. . . .	300,000	600,000
Total number of yards consumed	1,000,000	900,000	600,000	600,000
Price per yard to consumers	$1.00	$1.10	$1.50	$1.50
Cost to consumers	$1,000,000	$990,000	$900,000	$900,000
Additional cost because of tariff	. . . . .	90,000	300,000	300,000
Revenue to government	. . . . .	90,000	150,000	. . . .
Additional cost in excess of revenue	. . . . .	. . . .	150,000	300,000

Now suppose that next year a tariff is in effect at the rate of 10 per cent, which will make the cloth cost $1.10 in the home market. This price being still less than the home cost, there will be no home manufacture and all the cloth consumed will still be imported. In accordance with the normal demand curve, and assuming other circumstances unchanged, we assume that the amount consumed will fall off somewhat with the increase of price and that 900,000 yards will be taken in this year. The result is shown in column 2 of the table above. Consumers pay $990,000 for cloth which, were it not for the tariff, they could have obtained for $900,000, and this additional cost goes to the government in revenue. Here is obviously the case of a simple tax on imports.

Next year a 50 per cent tariff is in effect, and the imported cloth will be sold, if at all, at a price of $1.50. This price is just equal to the cost of making the cloth at home, and the foreign and domestic products are in competition on an equal basis. Presumably they will divide the market between them in some ratio, which for the sake of simplicity we assume will be an equal division. The result is shown in column 3 of the table, assuming a still further drop in the amount taken on account of the rise in price to $1.50.

Consumers now take 600,000 yards, paying for them $300,000 more than they would have paid for this quantity had there been no tariff. This additional cost however does not now all go to the government; the government collects 50 per cent upon the 300,000 yards imported, giving it a revenue of $150,000; but it gets no revenue from the 300,000 yards of domestic manufacture. We meet here for the first time the phenomenon of *protection*, by which we mean *encouragement given to a home industry by means of a tariff*.

Now let us proceed to the last case, that of the *prohibitive* tariff. A duty of 75 per cent being in effect, the foreign cloth can be sold only at a price of $1.75. Since the home industry can furnish the cloth (at least 600,000 yards according to our assumption) for $1.50, there will be no possibility of selling the imported article. This situation is shown in column 4 of the table. So far as the consumers are concerned, the situation is the same as in the previous year, except that all their purchases are of the domestic cloth. The government however now gets no share at all of the $300,000 additional cost imposed by the tariff upon the consumers. This is the extreme of protection.

Conclusions as to incidence and economic effects. To the reader who has fully grasped the meaning of the foregoing example, the most important principles of the customs tariff will now appear almost as a matter of course. They may be summarized in the following language (the reader will find it helpful to refer again to the numerical example above).

(1) A tariff will cause the price of the dutiable article to rise in the home market by about the full amount of the duty, up to a duty which is equal to the difference between the home and the foreign unit cost of production.[1] Above that the duty is prohibitive, and the further result depends on whether there is competition or monopoly in the home industry. If competition prevails, further increments of duty have no effect on the price; under monopoly

[1] The word "about" was introduced in this sentence in order to allow for the possibility of a change in the foreign cost of production resulting from the decreased purchase of the imported goods. The practical importance of this would generally be slight.

conditions, the price may rise by more than the difference in unit production costs up to the full amount of the duty.

(2) A tariff imposed at a rate less than the difference between the home and the foreign unit cost of production will have no immediate effect on home industry and will produce a revenue equal (disregarding costs of administration, etc.) to the added burden placed upon consumers.

(3) A tariff equal to or greater than the difference in unit production costs gives protection, and the additional burden placed upon consumers will be in excess of the revenue obtained by the government. If the tariff is in excess of the difference in unit production costs it is prohibitive, in which case the additional burden on consumers is accompanied by no revenue whatever to the government. Stated in slightly different words, any tariff which gives protection to domestic industry necessarily causes a loss to consumers greater than the gain obtained by the government.

(4) The additional burden placed upon consumers by a tariff which gives protection will be equal to the difference between the home and the foreign cost of producing all the goods (domestic and foreign) consumed, except when a prohibitive tariff accompanies a monopolistic condition in the home industry. In that case, the consumers may be burdened by an amount in excess of the difference in the entire cost of production.

Two purposes of the tariff. Tariffs may be imposed for either or both of two purposes: to obtain revenue or to give protection, the tariff being one of the most conspicuous cases of taxation used for purposes other than revenue. From our analysis thus far it is clear that these purposes are mutually antagonistic. Revenue comes only from the goods that are admitted; protection comes from the goods that are kept out. The extreme revenue tariff gives no protection; the extreme protective tariff gives no revenue. Nevertheless, in spite of the fact that revenue and protection work badly in double harness, they are teamed up in most of the world's tariff systems. Every nation in the world has a customs tariff for the purpose of revenue, and in most it is a very important source

of revenue. Virtually every country also makes use of the tariff in order to give encouragement to the development of industries within its borders, though there is almost infinite variety in the degree of protection thus granted or intended and in the classes of industries thus favored.

The revenue tariff. The traditional country of "free trade" is Great Britain. Since the reform of her tariff system in the middle of the nineteenth century, Great Britain has followed a policy of tariff for revenue with virtually no protection, although there has been some recrudescence of protective sentiment with a few protective duties in recent years, particularly during the World War. In order to make her tariff truly a revenue producer without disturbance from protective influence, Great Britain has adopted certain devices which are universally effective for this purpose. In the first place her duties are placed upon a comparatively small number of articles, these being generally commodities which in the nature of the case could not profitably be produced in Great Britain even in the face of high tariff duties. Such are tea, coffee, sugar, cocoa, spices, various fruits, and other tropical products. The second device for safeguarding the revenue tariff from defeat through the checking of imports and the development of home industries is to impose an internal excise duty on each home-made good to match the tariff duty on the imported article. This keeps the balance between the domestic and foreign products and prevents the tariff exerting any protective influence. Great Britain makes use of this device in certain cases. Finally there is the extreme safeguard of prohibiting outright the home industry, which is employed by Great Britain in the case of tobacco growing. By such means as these it is possible to build up a tariff system of extreme simplicity, easy to administer, having no disturbing influence upon the country's trade and industry, taking nothing from the people beyond what goes into the public treasury, and withal yielding a handsome revenue.

The protective tariff. Few countries however outside of Great Britain have chosen this kind of tariff. Virtually everywhere protection is consciously sought to a greater or less degree, and the

employment of devices to neutralize the protective tendency of duties imposed for the sake of revenue is rare. At the extreme protective pole stands the United States, where, in spite of the great importance of the tariff as a source of revenue, the protective purpose has always been strong or even predominant. The result is a tariff which, in striking contrast with the simple British customs, imposes duties at hundreds of rates upon almost every conceivable commodity and whose multiplicity and complexity of detail are such as to be intelligible only to the expert. Into these technical details it is fortunately not necessary for us to enter, our purpose being served when acquaintance has been made with the fundamental principles involved.

Fiscal weakness of the protective tariff. It has been shown that the two purposes of the tariff, revenue and protection, are not harmonious. Any nation which wishes to eliminate the protective feature entirely may make of its tariff a simple, efficient, and productive source of revenue as has been done by Great Britain. On the other hand a country which sought protection only and cared not for revenue could make its tariff serve well its purpose by imposing prohibitive duties on a selected list of commodities or, simpler still, by forbidding outright the import of such commodities. When however the attempt is made to combine both purposes in a broad way, as has been done by the United States, the effect is the best possible illustration of the principle, laid down in a previous chapter, that use of a tax for purposes other than revenue is apt to impair the tax as a revenue instrument.

The United States tariff is an instructive example. In certain years the tariff has produced revenues greatly in excess of the government's needs, while in other years dependence on it has caused revenues to fall far short of expenditures. During the first decade of the nineteenth century there were considerable surpluses in every year but one, the tariff being practically the only national tax. In 1807 the customs revenue was nearly double the year's expenditures. In 1837 the customs, still the sole tax resource of the national government, produced less than one third of the ex-

penditures and caused a deficit of over twelve millions, and deficits followed in every year except one down to 1843, in which year the customs yielded less than two thirds of the money needed to run the government. Two years from then the customs brought forth revenue well above the total expenditures, only to fall down again two years later (1847), when it was unable to produce half the amount of the year's expenditures. We have already noted how dismally the tariff failed to meet the war emergencies of 1812 and 1861. Since the Civil War the steadying influence of the excise taxes has somewhat mitigated the wild gyrations of the tariff revenue, but throughout our entire national history dependence upon the tariff has been, more than anything else, the cause of those startling successions of uncomfortable deficits and embarrassing surpluses which testify to an imperfect financial system.

The explanation is simple. The tariff being looked upon as an instrument for giving protection to certain industries, its rates cannot be raised or lowered to meet fiscal conditions without opening up the whole political conflict over protection. Furthermore the relation between tariff rates and tariff revenue is extremely complex, and it is impossible to predict accurately the effect of a change. By checking the flow of imports an increase of rates may diminish revenue instead of increasing it. A reduction of rates may increase revenue. In one political campaign, in which an unwieldy surplus revenue was an issue, both Democrats and Republicans promised to reduce the surplus, the one by lowering the tariff rates, the other by raising them; and both had good reason!

The controversy over protection. In spite of the widespread use of the customs tariff to give protection to home industries, the wisdom of this policy has for generations been vigorously challenged. Particularly in the United States has the dispute over "free trade versus protection" raged almost without interruption from colonial days to the present. This question is primarily an economic one, and it behooves us first to obtain a clear notion of the exact point at issue and then to seek the answer.

The principles of the shifting and incidence of import duties, as developed in the preceding pages of this chapter, and the principles

of interregional trade set forth in earlier chapters furnish us our starting point.[1] We know that, in the absence of restrictions upon its foreign trade, the people of each nation tend to buy goods in the cheapest market and sell in the dearest market and that this leads them to devote their productive resources and energies to those lines in which their nation has comparative advantage, as evidenced by the ability to produce at comparatively low cost. At the same time they rely upon other nations for those goods in whose production they are at comparative disadvantage. Thus develops international trade and international division of labor, with economic consequences which need not be rehearsed here.

International trade may be checked by means of duties on imports.[2] As we have seen, a duty high enough to be protective may diminish or entirely stop the import of any chosen commodity. Carried to its extreme limit, the protective tariff might be used to prevent virtually all imports. Even if applied only to those goods which are capable of production in the home country, the result would not be much different, since, barring only those commodities whose materials are not to be found within the territory, there is virtually no commodity which could not conceivably be produced if the foreign product were kept out. Doubtless such exotic products as tea and coffee would be produced in the United States — of inferior quality and at fabulous cost to be sure — if a high enough tariff excluded the imported articles. Whether applied moderately or to the limit, the protective tariff has the effect of reducing imports and building up corresponding industries at home. It thus restricts international trade, promotes " diversity of industry " in the home country, and diminishes reliance upon the goods of other nations. National self-sufficiency gains at the expense of international division of labor.

As to the truth of the foregoing statements no question can be entertained by anyone who is familiar with the elementary principles of economics. The controversy over protection starts from

[1] Chapters XXVI and XXVIII.

[2] Since they seldom enter into the controversy — never in the United States — we need give no consideration in this discussion to duties on exports.

this point. Granting the respective effects of freedom of trade and protection as we have stated them, the "protectionist" asserts that it is to the economic advantage of a nation thus to increase diversity of industry and national self-sufficiency; the "free-trader" [1] says that this means an economic loss and that the nation gains most when it enjoys all the advantages of international division of labor. Thus the issue is sharply drawn.

The case for protection. The advantages of trade, as already developed, are so evident and so important that we are justified in placing the burden of proof upon the protectionist, who claims that there is an economic gain from putting artificial restrictions upon a nation's foreign trade. Let us proceed then to examine the principal arguments advanced in support of the policy of protection.

(1) **The favorable balance of trade argument.** Some of the most popular arguments for protection are such naïve applications of old Mercantilist fallacies that they may be dismissed with nothing more than brief reference to the established principles of trade. Such is the argument that a nation gains in its foreign trade only from its exports, that imports are a national loss, and that the tariff is therefore beneficial through its effect in checking imports. We have already learned[2] that, so far from exports being superior, it is imports which are the real goal of international trade; we have learned also that a country's balance of payments must in the long run be in substantial equilibrium and that therefore anything which reduces imports must eventually also reduce exports. It is neither advantageous nor possible by means of a tariff to make imports decline relatively to exports.

(2) **The gold argument.** Similarly naïve is the argument that a favorable balance of trade, created by the tariff, causes gold to flow into the country and is therefore advantageous. This is sim-

[1] The reader should perhaps be reminded that "free trade," as the term is used, does not mean no tariff; it means no protection. The free-trader has no quarrel with the revenue tariff; he is opposed only to protection. There is no country in the world entirely without a tariff and no appreciable sentiment in favor of the abolition of all tariffs.

[2] See Chapters XXVI and XXVIII.

ply a corollary of the preceding argument and rests upon other Mercantilist fallacies with which the reader is acquainted. We know that it is not the balance of trade (*i.e.*, merchandise exports and imports), but the balance of payments (including all debits and credits) which determines the inflow or outflow of gold; and we are also aware that it is neither possible nor advantageous for a country to pile up gold as the result of artificial restrictions upon its foreign trade.

Gold and the national wealth. But there are those who are not content with this summary disposal of cherished beliefs. Is it not true, they say, that the rule for amassing wealth is to sell much and buy little? Is it not thus that the individual business man accumulates money from the profits of his business? And can we avoid the conclusion that the same rule must hold for national gain from trade? At the risk of adding proof of the obvious, we may give some further attention to these Mercantilist notions.

The wealth of the people of the United States is estimated (in 1922) at about 320 billion dollars. At that time the nation's total stock of gold coin and bullion was about 3,900 millions, or about $1\frac{1}{5}$ per cent of the national wealth. In 1880 the national wealth was, according to estimates of the United States Census, about 44 billions. The stock of gold coin and bullion was then 352 millions. In this period of 42 years, therefore, the national wealth has increased by 276 billions, while we have added but $3\frac{1}{2}$ billions to our stock of gold. It is certain enough that our gain in national wealth has not been due to accumulation of gold from foreign trade. The growth of a people's wealth is not principally in the form of gold; it consists in the increased value of its lands and houses, its factories and railroads and ships, its homes, and countless other forms of wealth infinitely more important than its stock of gold. And the gain from its foreign trade is not in the form of gold but in the useful commodities imported.

But what of the analogy of the business man? The answer is that neither does he accumulate wealth in the form of money or by selling more than he buys. If, at the end of a profitable year in which a man has made gains of $50,000, we should ask him to show

us his profits in concrete form, he would not point to an increase of his money on hand or his bank deposits. The probability is that these items are no larger than at the beginning of the year. His profits were not made by selling more than he bought. As he received money from the sale of goods, he spent it — partly for current consumption, partly for additions to his factory and machinery, for lands, for improvements to his home, for a new automobile, for government bonds and corporation securities. Thus he will show concrete evidence of his year's accumulation. So it is with the nation. "A nation, like an individual, grows wealthy, not by selling more than it buys, but by producing more than it consumes."

(3) **The added industry argument.** It is argued that the new industries which admittedly may result from protection are an addition to the industrial resources of the nation and that therefore the tariff has increased the national wealth. Let the reader recall that the tariff causes the establishment of a new industry by making it too costly to buy a foreign product previously imported. He will further recall that this diminution of our imports must be balanced eventually by a corresponding decline of our exports. The balance of payments must maintain its equilibrium; if we will not buy of foreigners they cannot buy of us. The result is that, while a new industry is arising, some old industry finds its foreign market declining and is forced to reduce its operations accordingly. The new industry is thus not an addition to the nation's industry but merely a diversion of capital and labor from one field to another.

It is true that, when a country is growing in population and enterprise, the diversion of industry induced by protection may take place without causing an absolute decline in any field. What the tariff does in this case is to direct into the protected industry new capital and labor which would otherwise have gone elsewhere. The development of the unprotected industry is stopped or diminished.

In either case, whether capital and labor already employed are shifted or new capital and labor are directed into the protected field, there is clearly no gain in total industry. Moreover this

diversion of industry is from the more profitable into the less profitable fields. In the absence of restrictive legislation, the nation's capital and labor are employed in those lines in which it has the greatest comparative advantage, and it is into these same lines that new capital and labor seeking employment in a growing country will come. The protective tariff prevents this natural disposition, forces industry into less advantageous lines, and so causes a net economic loss.

(4) **The wages argument.** This argument appears in two forms. (*a*) It is asserted that the protective tariff, by increasing the total amount of industry in the country, adds to the demand for labor and so tends to increase wages. Protection is therefore of advantage, not merely to the entrepreneur class, but to all laborers, and so to at least the great majority of the people. It is evident that this argument rests for its foundation on the added industry argument and falls with it. The mere diversion of industry, particularly when it is from more advantageous to less advantageous lines, can at best do nothing to increase the demand for labor or to raise wages.

(*b*) The wages argument is not always put so crudely as this. By American protectionists it is more often argued that, in the absence of the tariff to keep out cheap foreign goods, American manufacturers could not pay their present high scales of wages but would be compelled to lower wages or perhaps to go out of business entirely. Thus the tariff is in the interest of the workingman. This argument overlooks the fact that many American industries seem able to get on well enough without protection. These are the industries naturally fitted to prosper in the American environment, the industries in which America has a comparative advantage. Without protection, they are nevertheless able to pay the prevailing high level of American wages; they pay as high wages as the protected industries. If any industry is really unable to pay the prevailing rate of American wages it is because that industry is not adapted to our conditions. Its presence is due to the artificial stimulus of the tariff and it exists at the expense of some more favorable industry which could

pay high wages without tariff assistance. Removal of the tariff might cause the artificially favored industry to disappear, but its loss would be offset by the gain in the naturally favored industry, and there would be no decline in wages.

Wages and labor cost. There is much loose talk about wages in this connection. Protectionists argue as though, simply because wages are higher here than elsewhere, it must cost more to produce anything in America. They express alarm at the prospect of competition with the products of "foreign pauper labor." Now it is a well known fact that the wages of farm laborers in the Mississippi valley are higher than anywhere else in the world; yet there are few regions where the cost of producing wheat is so low. In numerous fields of manufacture goods are produced more cheaply in America than anywhere else in the world, in spite of high wages. The explanation of low production costs in such cases is of course our natural advantages of soil, climate, and mineral resources, our unsurpassed use of labor-saving machinery, and the extraordinary skill with which certain American industries have been organized. Under such conditions employers are able to pay high wages for high grade labor, and that without making the cost of production high.

American industry could not function on its present level of high efficiency if it had to do with only the foreign pauper labor to which the protectionist argument refers. Few American employers would be willing to exchange their present personnel of skilled labor for a complement of Chinese coolies even at the Chinese level of wages. The student will recognize that the countries whose competition our manufacturers most fear are not India, China, or Africa, but England, Germany, and France. Our most dangerous competitors are the countries of comparatively high wages. And the competition which European manufacturers fear includes America, the country of the highest wages of them all. In fact it is notable that among nations industrial efficiency goes with high wages, not with low wages.

In international competition the controlling factor is not high or low wages, but high or low *labor cost of production*. Wages, as such,

tell us nothing. It is only when wages are divided by quantity produced, to give the labor cost per unit of product, that we have the key to the problem of relative advantage in international competition. Thus it is that high relative wages when accompanied by still higher relative output give low labor cost. This is the condition in the great body of industries which are naturally adapted to American conditions. Here high wages are paid without the aid of protection. In general these are industries which either are favored by peculiar advantages of soil or climate or natural resources, such as cotton growing, wheat growing, copper mining, etc., or have the advantage of skillful use of labor saving machinery, such as the manufacture of automobiles, sewing machines, farm machinery, etc. Industries not so favored by nature or which must lean heavily upon hand labor do not flourish here. To them the high level of American wages is a real handicap, which can be overcome only by the protective tariff removing the competition of the cheaper foreign product. Protection is required therefore not to maintain the high level of American wages, but to make possible the payment of American wages in industries not suited to American conditions.

Historically, high wages in America antedated protection, and it was these very high wages that led to the demand for protection in lines which otherwise could not attract labor from those fields in which America offered the greatest rewards. In this sense it is correct to say that, so far from the tariff being the cause of high wages, high wages were the cause of the tariff.

Effect of withdrawing protection. The wages argument is without doubt urged in good faith by many American manufacturers who honestly believe that protection is necessary, not only to their own profits, but also to the well-being of American labor. Being engaged in lines of industry which possibly could not exist without tariff protection, they realize that without protection they could not continue giving employment to their labor forces at the present scale of wages. They visualize the distress which would follow a material cut in wages or a shut-down of industrial plants. They fail to see that discontinuance of the artificially fostered

industry would eventually open up other fields in which these laborers would find employment at American wages.

It is not to be denied that the shifting would impose inconvenience and loss upon the laborers and still more upon the capitalists. Industry that has become adapted to the artificial condition of protection cannot adjust itself to freedom of trade instantly or without some loss to all concerned. This is no argument for retaining permanently the artificial condition. But it does dictate caution and consideration in making the change. Protection has long been an established American institution; entrepreneurs have directed capital and labor into certain fields of industry relying upon the government's policy of protection. They and their employees have the right to ask that a change in policy be made gradually and with every possible care to minimize the unavoidable losses of adjustment.

The fact is that American industry would not be wrecked by a careful gradual removal of protection. Some few industries would doubtless succumb. In most cases certain grades or lines would be discontinued while others advanced, the sort of adjustment which now goes on in every progressive industry. For industry as a whole the final effect would without question be beneficial. American tariff history shows protection surprisingly less potent, either for good or for evil, than both its friends and its enemies would have us suppose. The course of American industry has not been vitally influenced by its imposition, nor would it now be revolutionized by its withdrawal.

(5) **The infant industry argument.** Protection has sometimes been urged as a means to facilitate and hasten transition from the agricultural to the industrial stage of national economy. A nation on the threshold of this change will be taking the first steps toward establishing numerous manufacturing industries. In these first steps manufacturing will have to contend with the competition of imported goods coming from older nations whose industries are already firmly established. They have the machinery, the technical knowledge, the skilled workers; they are controlled by experienced entrepreneurs. None of these essentials has as yet been

attained by the young industries of the first country, and the competition of foreign manufactures is felt as a severe handicap.

Under these circumstances the tariff may be used to give relief from foreign competition until such time as the young industries have become firmly established. This line of argument is known as "the infant industry argument" for protection. When logically used it rests upon the following assumptions: (1) the argument applies only to industries which are suited to the national conditions, (2) the obstacles to the establishment of such industries are temporary only, (3) the tariff protection need be temporary only, and (4) after the removal of the tariff the industries will be able to stand alone. When these conditions are fulfilled and when it is further proved that the nation will actually benefit from the establishment of the industries in question more rapidly than they would come in the natural order of events, the infant industry argument is impregnable.

Conditions in the United States at the beginning of the nineteenth century were fairly typical of those here assumed, and our first tariff acts were to some extent at least justified by this argument. Even more than the tariff, the international conflicts of the period from 1807 to 1815, by interfering with foreign trade, gave protection to the young American industries. However the conditions to which the infant industry argument was applicable had generally passed with the first third of the century. This argument is little heard of today. It must be noted also that it has not usually been urged in good faith; having used it to obtain tariff favors, its advocates have later shifted to other arguments to prevent removal of protection. Most protectionists today advocate permanent protection.

What is the territorial unit for protection? It is an interesting characteristic of the foregoing arguments for protection that each assumes that the tariff barrier will follow national boundary lines. Yet the arguments themselves are purely economic; they contain nothing to show that protection would not be exactly as beneficial if directed against a competing region within the same nation. We are tempted to ask why the New England textile mills should

not have protection against the rising industry of the South, and *vice versa*, or why the automobile industry in various states should not have protection against Detroit. Historically many of the first tariffs were imposed by towns or other minute political subdivisions. They proved such a burden on trade and such a drag on economic progress that they were swept away. In like manner it came to be recognized that the tariff systems of some score or more little states in central Europe were proving fatal to the economic development of that region, and the German *Zollverein* — customs union — swept them all away and opened the path to a later political union and to the marvellous industrial development of the German Empire. Originally the American states had each its own tariff system directed against the others. They had to give way in the interest of the economic welfare of the whole region. No one today would seriously propose to retrace any of these steps already taken in the direction of freedom of interregional trade. No argument exists to show that still further progress would not follow the removal of the barriers to all trade between nations.

Conclusion. There is no escape from the conclusion that the economic case for protection is fallacious. No gain can be shown to offset the demonstrated advantages of international division of labor. Protection is an economic burden, a cause of loss, to the people of the nation which employs it. The economic advantage is all on the side of freedom of international trade.

Military necessity. Protection may be urged from an entirely different standpoint. The tariff is necessary, it is said, in order that the nation may be assured of a supply of certain products vital to its safety in time of war. If it be proved (1) that a certain industry is thus necessary and (2) that this industry will not exist unless relieved of foreign competition, then it follows unquestionably that protection is advantageous. Whether these two assumptions are actually fulfilled is a matter of fact to be determined in each case; it is doubtful whether, in a country of the size and variety of resources and industries of the United States, there is much room for this argument. It should be noted that this argument for protection is not based upon economic grounds. There is no

false assumption that the people will be richer on account of the tariff. It is admitted that they will be poorer, through having to pay more for some important commodity than they need have. But so are they made poorer by the taxes which defray the cost of building battleships. The justification is the same in either case.

Protection by bounties. Occasionally the object of protection is attained, not by a duty on the foreign product, but by a bounty paid by the government to the domestic producers. The United States for example has on occasions sought to encourage the growing of sugar by means of bounties. The economic significance of the bounty is the same as of the protective tariff, with one important exception. The cost of protection by tariff is imposed upon the consumers in proportion to their use of the goods in question. The cost of the bounty rests upon the whole group of taxpayers. Since protection is to be justified only on the grounds of public advantage, the apportionment of the cost under the system of bounties is more equitable than that of the protective tariff. Nevertheless bounties are very rarely used, probably because it is here so evident that protection costs something, that a burden is placed on the whole people for the benefit — immediate at least — of a favored few.

Ship subsidies. Ship subsidies are bounties paid to ship owners or operators in order to encourage the development of national shipping — the "merchant marine." The economic nature of ship subsidies is precisely the same as of protective tariffs and bounties. The justification of subsidies to shipping must be found, if at all, in the military necessity argument.

CHAPTER XLIX

OTHER FORMS OF TAXATION

Excise taxes in the United States. An excise is in general a tax levied upon the manufacture or sale or consumption of commodities or the exercise of an occupation or enjoyment of a privilege within the limits of the jurisdiction imposing the tax.[1] What will chiefly interest us is the taxes levied upon the manufacture and sale of commodities in the country, such as are commonly known in the United States as " internal revenue " taxes.

We have already given cursory notice to the part played by excise taxes in the history of the United States revenue system up to the time of the World War. That conflict put unprecedented demands upon the excise system beginning as early as 1914, since the disordered conditions in Europe weakened our revenue system for three years before our own entry into the war. Rates were increased on beer and wines; special taxes were imposed on bankers, brokers, theatres, etc.; railroad and Pullman tickets, freight transportation, telegraph and telephone messages, theatre tickets, and club dues were taxed, soft drinks were subjected to taxation hitherto confined to alcoholic liquors; and a large group of new commodities was added to the taxable list. By such means the revenue from the excise system was with little delay about trebled, averaging 875 millions during the five years 1916–1920 and reaching about a billion and a quarter in 1920. In the revenue acts of 1921, 1924 and 1926 the greater part of these taxes has been repealed, though several still remain as reminders of the war. On the other hand the adoption of national prohibition has almost done away with

[1] This definition, the one generally accepted by the authorities on public finance, is adopted for use in this book with full recognition of the fact that the term has no universally accepted meaning, but is used, particularly in judicial decisions and legal writing, with a variety of meanings, some very much broader than the one here chosen.

what was formerly the most important part of the peace-time excise system. In spite of these reductions the excise taxes are still far above the pre-war level of yield and rank next to the income tax among the tax resources of the national government, having definitely and probably permanently forged ahead of the tariff.

Shifting and incidence. The excises present the problem of shifting and incidence in its simplest form. Consider for example the case of the United States internal revenue tax upon cigars. The tax is imposed upon all manufacturers of cigars at a certain rate per hundred and is thus an addition to the normal unit cost of production. Manufacturers are therefore able to increase the selling price by practically the amount of the tax and so to pass the burden on to the wholesale dealers to whom they sell their product. These in turn, through the influence of their costs upon their selling prices, pass the tax on to the retailers who likewise pass it on in the prices charged the consumers. Sometimes the tax may be originally imposed upon the dealers, wholesale or retail, instead of upon the manufacturers. In this case the result is the same, though the process of shifting has fewer steps before the final consumer is reached. Again the tax may be imposed upon the consumer himself at the time of purchase of the taxed article, the law requiring the dealer to collect the tax and pass it back to the government, as was done with the war tax on soda fountain drinks, theatre tickets, etc. In such a case the price charged by the dealer is of course increased by the amount of the tax. When the excise is based upon an industrial occupation, instead of upon a commodity, the same result follows; the tax becomes an added cost of conducting the particular occupation and is passed on to the consumers by means of a higher price for goods or services. The former saloon-keeper's license was a familiar example; another example is the special war taxes on proprietors of theatres, circuses, and other commercial amusement places. The general principle is that the incidence of excise taxes on commodities or occupations is upon the consumers of the goods or services involved, the taxes being shifted when not imposed immediately upon the consumers. The excises, like the customs, are therefore commonly called "consumption" taxes.

Effect on production. It is not to be inferred that an excise tax is wholly without burden upon the manufacturers and dealers; indeed the reverse is generally true, due to the normal shape of the curves of demand and supply. Let *dd* and *ss* (Figure 8) represent respectively the demand and supply curves of a certain commodity, determining the selling price *pa*. Now suppose an excise tax is imposed upon the manufacturers or dealers. The unit cost being thereby increased, there will be a new supply curve *s's'*, higher than the former curve. The intersection of this curve with the demand curve, of course not changed by the tax, will determine a new price *p'b*. This price is higher than the old price, but it is frequently not increased by the full amount of the tax and, what is of more importance, the quantity sold has fallen off (from *oa* to *ob*). The manufacturers and dealers will, after the process of shifting, be able to charge enough to give them the ordinary rate of profits but only on a smaller volume of business.

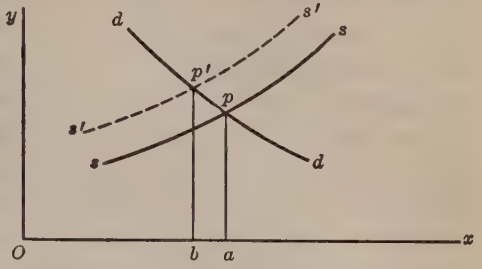

FIG. 8.

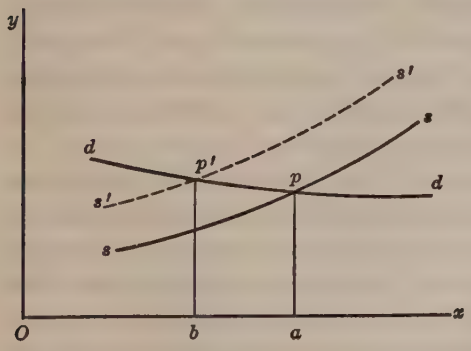

FIG. 9.

Elasticity of purchase. The amount of this loss to producers depends upon the degree of elasticity of the purchase. If purchase is extremely elastic, (see Figure 9) the increase in price occasioned by the tax may cause a sharp decline in consumption, thus causing great curtailment of production and driving many producers out of business. This gener-

ally occurs when some satisfactory substitute is left untaxed. On the other hand, if purchase is inelastic (see Figure 10), a tax may have little effect upon consumption and the volume of production. In such cases the consumers make the necessary saving in some other line of consumption. For example it is doubtful if the war tax on admissions had a great effect upon attendance on the "movies."

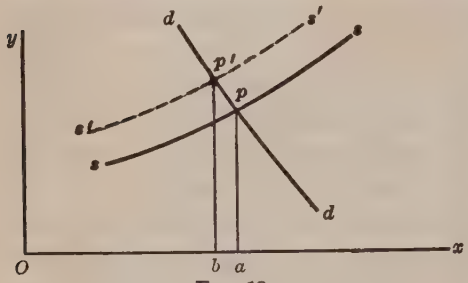

FIG. 10.

Technical and administrative features. The revenue purpose of excise taxes, in contrast with the tariff, is seldom much complicated by other motives. In the aim to get the maximum of revenue with the minimum of administrative labor and expense, full account is usually taken of the principle of shifting and incidence outlined above. Thus the articles selected for taxation are usually those of wide general use, if not "necessities" at least articles with inelastic purchase. This is in order that the consumption may not be greatly reduced by the imposition of the tax; otherwise the revenue would fail. This principle explains the almost universal taxation of liquors and tobacco; it explains the burdensome salt taxes imposed by some nations. Excise taxes are often imposed upon "luxuries," but if the taxed articles are really luxuries used only by the wealthy or well-to-do, no very large revenue is to be expected, because of the relatively small amount of consumption.

The desire for administrative simplicity generally dictates the selection of a short list of articles capable of bearing heavy taxes, rather than a comprehensive system of excise duties, a principle which may also be carried out in a revenue tariff system. The same consideration generally leads to the imposition of the tax upon the manufacturers or dealers rather than directly upon the consumers, since in this way the number of the immediate taxpayers is reduced to the minimum. A good many of the excise taxes used by the United States in financing the World War, such as those on

soda fountain drinks, telephone and telegraph messages, and the so-called "luxury taxes," were collected by the dealers directly from the consumers. The result was much annoyance to the public, trouble and expense to the dealers, loss of revenue through evasion, and general dissatisfaction leading to the term "nuisance taxes" with reference to many of these excises. The ordinary rule of the United States internal revenue is the opposite one of approaching the few manufacturers rather than the multitude of consumers. Both specific and *ad valorem* rates are used for excise duties, though the former are more favored on account of their greater ease of administration. A very effective administrative device is the use of adhesive stamps, such as adorn every box of cigars or package of cigarettes or pack of playing cards in the United States; this makes evasion difficult and collection easy.

Consumption taxes generally regressive. Consumption taxes, whether excises or customs, are generally regressive. The articles of wide use and inelastic purchase upon which they mostly rely call for a larger part of the total expenditures of the poor than of the rich, and their burden therefore becomes lighter as the taxpayer's income increases. A tax system based exclusively or principally upon excises and customs therefore runs counter to the generally accepted principle of progressive taxation and bears with great severity upon the poorer classes. On the other hand the principal direct taxes, income tax and property tax, generally reach only the more prosperous classes. It is only by means of the consumption taxes that all the citizens may be called upon to bear some share of the cost of government. Most tax systems are therefore, like that of the United States, a combination of direct taxes on property and income with customs and excise taxes, along with certain less important taxes which we shall presently study.

State excise taxes. Excise taxes are generally used by national rather than by local governments. This is principally because of the tendency of such taxes to drive business outside the borders of the taxing jurisdiction. A local excise on the manufacture or sale of a commodity would tend to send the manufacturers and dealers into a neighboring town or county or state where no such tax

existed and from which they could send their products freely to their former customers. A local tax has to be on something like real estate which cannot migrate. A national government does not have this fear; most industries supplying its citizens must remain somewhere in the country, and even if it were otherwise practicable to move to a foreign country and send in the products this attempt at escape will be readily met by imposing customs duties at least as high as the excises. In the United States the several states have generally been deterred from extensive use of excise taxes by these considerations, enforced by the constitutional prohibition of customs tariffs and excises burdening interstate commerce. There is still possibility of some greater use of excise taxes, as is evidenced by the recent Pennsylvania tax on coal and particularly by the phenomenal development in the past few years of state taxes on the sale of gasoline. In general however the states will always be very careful to impose no tax that may frighten away business.

Taxation of corporations. At first sight it may possibly be hard to see why the taxation of corporations should present a topic for special consideration. As a legal person, should not the corporation be amenable to all the ordinary forms of taxation exactly like a natural person? As regards such taxes as the tariff and the ordinary excise, there is indeed no reason for any distinction; an importer is an importer, whether incorporated or a natural person, and so of a manufacturer or dealer or consumer subject to the excise taxes. In the field of property and income taxes however the entry of the corporation has raised some of our most interesting tax problems.

Corporations and the income tax. The income tax, as we have already had occasion to observe, is essentially a personal tax. Its purpose is most nearly attained when the definition of taxable personal income is made as broad as possible. Such broad definition would include all income, in the form of dividends and interest, arising from investment in corporations. In such case it is obvious that the income of a corporation does not represent an additional class of taxable income. The only permissible income tax upon corporations would be a tax upon that part of their income

which was not distributed to their investors in dividends and interest; anything beyond this would be double taxation. These are principles which are generally recognized and not often consciously violated. What usually brings the corporation within the purview of the income tax is the attempt to apply the device of stoppage at source or to accomplish a similar purpose. Thus the British income tax requires the corporation to deduct the basic income tax from the dividends and interest paid to its security holders. This is in no sense a tax upon the corporation; it is merely the requirement that the corporation shall act as tax-gatherer for the government and is real stoppage at source. A *quasi* stoppage at source is attempted when the corporation is made itself subject to the income tax in lieu of taxing the stockholders on dividend income. Since any tax upon the corporation reduces by so much the funds available for distribution to the stockholders, why not give up the troublesome task of levying an income tax upon each individual stockholder and obtain the same result by one tax upon the corporation's net income? But the same result will not be obtained except when the income tax is levied at a uniform flat rate and when no personal exemptions of any sort are allowed. The impossibility of such adjustment in connection with an income tax involving progressive rates and personal exemptions has been demonstrated by our previous study of the United States income tax.[1] Equity, if not simplicity, is always sacrificed when there is departure from the strictly personal nature of the income tax.

Difficulties with the property tax. The most important problems of corporate taxation arise in connection with the general property tax. When corporations made their initial appearance in America in the first half of the nineteenth century it was assumed that, being legal persons and owning property, they would be subject to the general property tax exactly like other property owners. This was in fact the rule, no new legislation being ordinarily required to accomplish it. At the same time the stocks and bonds of corporations were taxable personal property in the hands of their owners. Thus there arose that form of double taxation to which

[1] See Chapter XLVII.

reference was made in Chapter XLVI. Since it was obviously easier to tax the corporation once for all than to seek out and tax each individual security owner, the taxation of corporations generally continued. Some states sought to avoid double taxation by giving up the taxation of stocks and sometimes also of bonds but, since the property tax is a local tax and since the state or town where the corporation's property is located is not necessarily the same as that of the residence of the stockholders and bondholders, this remedy was capable of only partial and imperfect application. In many states double taxation has continued with little or no relief.

Soon there arose other and more serious difficulties. It is in connection with the assessment of the property of the great corporations that the hopelessness of the task imposed upon the assessors becomes especially apparent. The local assessor could not cope with the problem of assessing the complicated machinery and materials of a great manufacturing corporation, still less the property of the railroads, the telephone and telegraph corporations, the banks, the insurance companies, etc. Furthermore the wealth of many of the great corporations is spread over many towns and counties and even into neighboring states. What could be more ridiculous than the attempt of the local assessor to value a few miles of railroad track or telegraph wires or a few hundred telephone poles? There also appeared to be something wrong when the heavy taxes on valuable terminal property, such as stations, roundhouses, and shops, or on great office buildings, warehouses, etc., all went to a few lucky towns or counties in which they happened to be situated, while other towns had to be content with taxation of a few miles of right-of-way and tracks. And finally there was the problem of assessing locomotives and railway cars which passed daily through various towns but had no permanent location.

The technical difficulty of valuing complex properties has thus far not been enough realized to lead to any great improvement. As we have already seen, the local assessors throughout the greater part of the United States are still struggling almost unaided with the problem of assessing the complicated buildings, machinery, and goods of manufacturing and mercantile corporations.

Public utility corporations. As regards the great public utility corporations — railroads, express companies, car companies, telephone and telegraph companies — the situation became so intolerable that reforms, more or less successful, have been instituted in most of the more progressive states. These reforms fall generally into two classes. In a number of states the assessment of the property of such corporations has been taken from the local authorities and intrusted to state boards or commissions which are thus enabled to value the entire property of each such corporation (or at least as much of it as is located within the state) as a unit. The value thus determined is then generally distributed among the several towns and counties, which are permitted to tax each its share at its own tax rate.

Experience has generally shown however that the property basis is poorly adapted to the taxation of the great public utility corporations and many of the most advanced states have therefore given up the property tax and substituted other methods for taxing their corporations. The most common method is the tax upon the gross earnings from operation; less frequently the tax is on net earnings. A system which is now gaining favor is the " gross-net tax," a compromise method by which the tax is imposed upon the gross earnings at a rate which varies with the ratio of net earnings to gross. Another common method is to place a tax upon the value of the shares of stock or upon the value of shares and bonds, known as the " capitalization " basis. In some of these cases the departure from the property basis is not complete, in that the local bodies retain the right to tax the corporation's real estate wherever it is located.

Taxation of financial corporations. Considerable reforms have also been achieved in the taxation of the great financial corporations, such as banks, trust companies, savings banks, insurance companies, etc. Banks are usually taxed upon the value of their stock, and trust companies are frequently treated similarly. Savings banks are commonly taxed upon their deposits, insurance companies upon premiums collected, etc. These taxes are usually levied by the state, often with distribution of the revenue to the

towns. In most of these cases, the real estate of the corporations is subject to local taxation under the general property tax.

Exemption of securities. When corporations are taxed upon all their property or are taxed by other methods equivalent to taxation of their property, double taxation will be avoided only if the stockholders are exempt from taxation upon the shares of stock. There is on the whole adherence to this rule in most of the states in which the taxation of corporations has been seriously overhauled. Elsewhere double taxation of this sort is still common. As to the bonds, exemption of the owners is not so simple a remedy nor is it so commonly granted. The whole adjustment is made difficult and unsatisfactory on account of interstate complications arising when the security owners are residents of states other than that in which the corporation or its property is located.

Special excise taxes on corporations. The special treatment of corporations is usually closely related to the general property tax, in that it consists either in a modification of the property tax or in the substitution of special forms of taxation in lieu of the property tax. There are however examples of special corporation taxes not related to the general property tax. These are usually to be regarded as special business taxes or excise taxes upon the privilege of doing business in the corporate form. For example Connecticut thus imposes a state tax upon the net income of manufacturing and mercantile corporations in addition to the local taxes upon their property.

Business taxes. Taxes imposed with respect to the conduct of business are not uncommon, though of less importance than those forms of taxation which have thus far been studied. Reference has just been made to certain special taxes upon corporations which fall into this classification. License taxes upon individuals and corporations engaging in certain specified occupations are common in the Southern states of the United States and are often quite far reaching and productive of considerable revenue. Such taxes are not important in the more advanced industrial states, and their significance is likely to decline as time goes on. Before the days of prohibition the state and local license taxes upon retail liquor

dealers were a well known example of the business tax. There are various examples of other local license taxes, such as those on peddlers, hack drivers and taxicab operators, auctioneers, etc., some of which may however be more properly regarded as fees. On the whole the special field of business taxation is not very deeply cultivated in the United States, though there are some indications of an increasing interest.

The inheritance tax. There is sooner or later one occasion upon which all property of every description must be transferred, namely upon the death of its owner. Even in the case of property owned by corporations, whose lives are of indefinite duration, securities representing rights against all the property are owned by natural persons and must be transferred upon their death. This gives to government a convenient opportunity to impose taxes from time to time upon the transfer of property of every description, an opportunity which is availed of by sovereign governments quite generally the world over. Such taxes are called inheritance taxes, succession taxes, estate taxes, death duties, etc. They are actually a form of excise taxation, both economically and legally, but they present peculiarities which make it convenient to study them thus by themselves.

Typical technical features. The typical modern inheritance tax contains the following technical features. At the outset property left to certain religious, educational, or charitable institutions is generally entirely exempt from the tax. Of the remaining taxable estate, a certain amount is usually exempt, this amount varying according to the degree of relationship between the decedent and the beneficiaries. Thus a considerable value may be transferred free of tax to wife or husband, children or parents; smaller exemptions are allowed in diminishing scale as the degree of relationship becomes more remote till the smallest exemption, or sometimes no exemption at all, is granted to the most remote relations and strangers. Above the exempt minimum, the tax is calculated at rates which are usually doubly progressive: (1) the rates start low for bequests to the closest relations and increase step by step as the degree of relationship becomes more remote, and

(2) within each class of beneficiaries, the rate is progressive according to the amount transferred.

As an illustration of the typical plan of inheritance tax rates we may take the following classification of the Connecticut successions tax.[1]

	Class A	*Class B*	*Class C*
Exemption,[2]	$10,000	$3,000	$500
Above the exemption, up to and including $25,000	1%	2%	5%
Above $25,000 up to and including $100,000	2%	3%	6%
Above $100,000 up to and including $200,000	3%	4%	7%
Above $200,000	4%	5%	8%

There is of course opportunity for indefinite variety in the details of classification, exemptions, and rates. Among the states of the United States the top rate on direct heirs varies all the way from 3 per cent to 14 per cent; on the most remote relations and strangers the maximum rates run as high as 40 per cent.

Sometimes the inheritance tax is levied upon the entire estate without relation to the amounts of the distributive shares or the relationship of the beneficiaries; such a law is more properly called an estate tax. As an example we may take the United States estate tax, as contained in the revenue act of 1926. The tax, as regards resident decedents, is imposed only upon estates whose amount exceeds $100,000 and is based upon the "net amount" of the entire estate, calculated by subtracting a lump sum of $100,000 besides other deductions. With respect to the net estate so defined the rate begins at 1 per cent upon the first $50,000 and increases progressively upon the successive increments of the net estate till the maximum rate of 20 per cent is reached, applying to that part of the net estate in excess of $10,000,000.

Inheritance taxation by the American states. The inheritance tax, although a comparatively modern institution, is today a

[1] Class A includes husband and wife and ancestors and descendants in the direct line including those adopted; Class B includes husbands and wives of children, step-children, and brothers and sisters and their descendants; Class C includes all others.

[2] One exemption only in each class. In most states, the exemption is allowed for each beneficiary in the class.

recognized part of the tax system of practically every civilized nation of the world. In the United States the attention of the states was first directed to inheritance taxation little more than a generation ago. Since then the development has been rapid and widespread. State after state has fallen into line until today only a few states fail to make some use of this resource.[1] At the same time there has been a continuous movement to broaden the tax by extending it to new classes of heirs and by lowering exemptions, increasing rates, steepening the progression, refining administrative procedure, etc. The revenue thus obtained by the states has increased from an insignificant amount forty years ago to over $8,000,000 in 1902–1903 and nearly $75,000,000 in 1923. Among several of the more populous and wealthy states, inheritance taxes regularly constitute from five to fifteen per cent or more of the total state revenue.

The federal estate tax. The Federal Government of the United States has inclined to look upon the inheritance tax as an extraordinary source of revenue for war emergency. A federal estate tax was in effect from 1798 to 1802. After that the next use of this tax was from 1862 to 1870 on account of the Civil War. The Spanish War revived the tax during the years 1898 to 1902, and finally the World War led to the imposition of the estate tax which still continues, and from which the government obtains an annual revenue of something over $100,000,000, about 3 per cent of the total ordinary receipts.

State and inter-state complications. Because of our federal form of government, the inheritance tax presents a peculiar problem in the United States. Each of the states is a law unto itself, and there is consequently the utmost variety in the exemptions, rates, and other details of our more than forty state inheritance taxes. It is natural, unfortunately, that each state should incline to obtain all the revenue possible from its tax without much regard to the interests of the other states or their citizens. There has resulted a chaotic condition of overlapping jurisdictions and mul-

[1] At the close of the year 1925 only three states, Alabama, Florida, and Nevada, and the District of Columbia were not imposing some form of death tax.

tiplicity of tax obligation which almost beggars description. There is little difficulty with real estate, whose transfer can as a rule be taxed only by the state in which it is located. Tangible personal property, when located outside the state of the late owner's residence, has commonly been taxed both by the state of domicile of the decedent and by the state in which it was located at the time of death. This part of the problem appears now to have been solved by a decision of the United States Supreme Court (in Frick *vs.* the Commonwealth of Pennsylvania, announced June 1, 1925) holding that tangible personal property, like real estate, may be taxed only by the state in which it is located.

As to intangible property the situation is confusion worse confounded. The state of the late owner's domicile will almost always claim jurisdiction over the whole amount; in addition, if the estate contains stocks or bonds of a corporation chartered by another state, that state may also impose its tax upon the transfer of such securities; if the corporation maintains an office for registry or transfer of its securities, the mere location of this office may give jurisdiction to a third state; and finally the physical presence of the securities in a safety box in another state may be the basis of a claim from a fourth state. A given estate may thus be subject to the taxes of any number of states, in addition to the federal estate tax.

This results in double and multiple taxation whose amount in some cases is very severe or even confiscatory. Furthermore the details of the several tax laws are so diverse and their administrative features so complicated that estates which are thus involved can not be settled without great sacrifice of time and labor, and the expense of administration thus often exceeds the total amount of the taxes finally found to be due. The seriousness of the situation thus created has rapidly increased during recent years, till it threatens to discredit this very important branch of taxation. Efforts are now under way by various organizations and conferences to find a solution, whereby taxation of inheritance shall be reasonably uniform throughout the United States and double and multiple taxation of the same property shall be avoided.[1]

[1] Cf. *Report of the National Committee on Inheritance Taxation*, 1925.

Some fundamental questions. There are certain questions of broad principle which arise in connection with the inheritance tax and which must receive at least our brief attention. As to its incidence, the inheritance tax is of course a burden upon the beneficiaries of the estate, though the effect upon the separate classes and individuals varies greatly according to the details of the law and the provisions of the particular will. A bequest, except where received by a dependent of the testator, comes as a more or less extraordinary windfall, different from one's ordinary regularly recurring income. This is usually taken to justify quite heavy taxes upon inheritance as compared with the other kinds of taxes on property, income, etc. Similar conclusion is drawn from the prevailing concept of property rights, which holds them to be distinctly less sacred when it comes to disposition of one's possessions after death, especially that part which is not passed on to those immediately dependent upon the testator. There are those who would go still further and use the inheritance tax as an instrument to check the rise or perpetuation of large family fortunes. There is without doubt considerable sentiment in favor of this "socialistic" use of the inheritance tax, but the sentiment has not thus far become strong enough or intelligent enough to exert any great practical influence. It is significant that in all inheritance taxes having progressive rates based upon the distributive shares, the progression goes from the closer to the more remote relationship, whereas the reverse would be the rule if curbing large family fortunes were the end chiefly sought. On the whole, so far as inheritance tax rates are not determined by mere rule of thumb, they represent a more or less fair application of the principle of distribution of taxation according to ability to pay.

Inheritance tax rates, at least as regards bequests to indirect heirs and strangers, might be pushed very high before reaching any limit set by the principle of ability to pay. But before this limit is reached there is almost certain to be a limit set either by administrative considerations or by broad social policy. There are all kinds of business arrangements and legal devices whereby inheritance taxes may be escaped while still keeping within the letter of

the law. There is always the possibility of escape through gifts before death, a door which no development of gift taxes is likely completely to close. The higher the rates the more will such avenues of escape be resorted to. The second limit is reached when the tax becomes so high as to discourage the accumulation and husbanding of capital or to interfere with its most productive use. While it is probable that this effect of the inheritance tax has been exaggerated, there would eventually come a time when the public loss from decline of capital and disturbance to business and investment would exceed the possible benefits from the revenue thus contributed.

The poll tax. This tax is mentioned mainly to avoid an apparent omission. It is a tax, generally at a uniform rate of one or two or three dollars, upon every adult person, or every voter, or some other broadly inclusive class of the people. It conforms to no recognized principle of tax distribution, except the desire to compel every person to make some contribution to the cost of government. It is expensive to collect, except from those who are already subject to the property or income tax, and while it is (on paper at least) a recognized part of the American local tax system, it is generally honored more in the breach than in the observance, and the revenue derived from it is usually ridiculously small. Nothing is to be hoped from reform of its administration and nothing to be gained from its perpetuation.

CHAPTER L

GOVERNMENT INDUSTRY

We have found it convenient to divide the activities of government into two groups, the primary functions and the industrial, and we have noted the sharp distinction between the methods used to finance these respective groups.[1] It is a commonplace that private enterprise is the rule in modern industry. Government industry is the exception, and we shall find that the government generally has some special purpose in view when it sees fit to invade the industrial field. We have already seen that this purpose is only rarely the ordinary business motive of profit-getting.

The reasons for government industry. Government ordinarily undertakes an industrial enterprise either because its control of the enterprise in question is considered essential to the performance of some of its primary functions or because it is felt that some essential commodity or service can thus be provided for the people more effectively or more equitably than by private industry. When the telegraph and telephone were first introduced, it was recognized by the governments of the leading nations of Europe that these instrumentalities were destined to be of great military importance, and it was felt that it would not be safe to leave them in private hands. From the beginning the telegraph and telephone have been government enterprises in most of the European nations. This consideration is also the chief explanation of the prevalence of government ownership of railroads in Europe. The United States, remote from the militaristic atmosphere of Europe, has not been moved by this motive, but it is interesting to note that at the time of the World War military considerations led the

[1] See Chapters X and XLIV.

United States to assume temporary control of the railroads and the telegraph and telephone services.

The Government Printing Office is an example in the United States of an industry undertaken to facilitate the performance of the government's primary functions. The most important and universal example is the post office. In most European countries the postal service also was originally undertaken by the government for military or other governmental purposes. But the principal reason for the government post office today is the political motive. Easy and cheap communication, open to all, tends to bind the people together with common ideas and ideals — to foster a "national spirit." Public opinion has opportunity to crystallize and to express itself, and the government may be informed of the wishes of the people. These ends are felt to be vital to the well-being of the modern democratic state, and it is feared that they would not be so fully accomplished if postal service were left in private hands. The forestry policy of the United States Government is dictated by somewhat similar considerations. Recognizing the vital importance of lumber and other wood products and foreseeing the decline of our natural sources of supply under private ownership, American public opinion has approved the acquisition and management by the Federal Government, and by some of the states as well, of a great public forest reserve.

A somewhat different motive appears when the government takes over an industry in order to do away with the abuses of private monopoly. Though quite common in Europe, few actual examples of this motive can be cited in America.

Government ownership in the United States. American soil has not been favorable to the development of government ownership. Barring a few enterprises set up to supply the government's own needs, such as the Government Printing Office, the Federal Government of the United States has only one great industrial enterprise, the post office.

The states have even less. In the first half of the nineteenth century several of the states ventured rather extensively into the industrial field, particularly in railroading and banking. These

state enterprises were almost without exception disastrous failures; the states lost heavily and in certain cases defaulted upon the bonds by which the capital investment had been obtained. The lesson was so thoroughly learned that to this day the states have been generally quite willing to keep out of industry.

Among the cities and towns of the United States, examples of public industries are somewhat more common. Municipal water supply systems are by no means rare, and there are a few municipal gas systems, street railways, and other examples. But even among the municipalities we do not find any great development of government ownership, and the movement does not at present appear likely to gain in significance.

The United States Post Office. As the one example in the United States of a great government industry, the United States Post Office is worthy of study, both for its own sake and because it may serve to exemplify some of the important principles of government industry in general. That the post office is a great industrial enterprise is sufficiently evidenced by its payroll of 300,000 employees and its annual receipts of over half a billion dollars.

Postal rates. Assuming that the more obvious features of our postal organization are known to everyone, let us proceed at once to a somewhat critical analysis of the scheme of postal rates, which is the most characteristic and fundamental feature of the business. Mail matter is classified into four classes, as follows:

First class. Letters and all other written matter, as well as all packages that are sealed. *Rate:* 2 cents per ounce (government postal cards, 1 cent; private mailing cards, 2 cents).

Second class. Newspapers, magazines, and other periodical publications. *Rate:* (*a*) When mailed by person other than the publisher or news agent, 2 cents for each 2 ounces, up to 8 ounces. Above 8 ounces, at fourth class rates. (*b*) When mailed by the publisher or news agent, that part of the publication which does not consist of advertising matter pays $1\frac{1}{2}$ cents a pound; on advertising matter the rate varies from 2 cents to 9 cents a pound according to the distance.

Third class. All other printed matter, such as books, pamphlets,

circulars, etc., and merchandise, but not exceeding 8 ounces in weight. *Rate:* 1½ cents for each 2 ounces.[1]

Fourth class (*"parcel post"*). All matter weighing more than 8 ounces, except first class and second class matter. *Rate:* Service charge of 2 cents per package plus a charge varying with weight and distance.

The flat rate. This scheme of rates has been presented thus in detail because it raises certain important questions of principle in connection with the postal business. We notice first that in the first and third classes and to a considerable extent also in the second class the rates charged are based upon weight alone without reference to the distance carried. This principle of the "flat rate" is a characteristic feature of the postal business — at least as regards letters — all over the world. It is justified on two different grounds. From the administrative point of view, by far the greater part of the cost of handling mail matter is caused by collecting and delivering at the two ends and has nothing to do with the distance carried. It becomes possible to average the carrying cost without undue arbitrariness and of course there is thereby much gain in simplicity.

The other and more fundamental basis of the flat rate charge is the political purpose to make the post office a means of uniting the whole people through cheap and widespread communication. While in general the cost of carrying is relatively insignificant, this is not true of the more remote and inaccessible places. Under the flat rate system, residents of those regions are not called upon to pay the very heavy costs of their own postal service but are given the advantage of the low average rate at the expense of the residents of the more thickly settled regions. The additional burden on the latter however, when spread over the whole business, is insignificant and is accepted without complaint in view of the important ends to be served.

The foregoing is true of letter postage. The balance of advantage and disadvantage swings the other way when the post office

[1] There is a special rate in this class of 1 cent for each 2 ounces on "books, catalogs, seeds, cuttings, bulbs, roots, scions, and plants."

goes extensively into the carrying of merchandise. Here the carrying cost, depending on distance, is by no means insignificant, and the flat rate cannot be justified. The flat rate was employed in all four classes of the United States postal service until the introduction of the parcel post with the "zone" system on January 1, 1913. It was not till 1917 that the principle of differentiation was extended to periodical advertising matter in the second class.

Rate discrimination. The most striking feature of the United States postal rate structure is the wide differentiation between the rates of the several classes of mail matter. Reduced to a pound basis, the primary rates of the first three classes are respectively 32 cents, 1½ cents (publishers' rate), and 12 cents (books 8 cents). Letters are charged more than twenty times the rate imposed upon periodicals, four times the rate on books, and nearly three times the rate on small merchandise packages; books are charged more than five times as much as periodicals. Of course the high pound rate for the first class is in part justified by the fact that the individual packages are so much smaller than in the other classes. Yet if we reduce the rate to a piece basis, the discrepancy is still very great. The average revenue per piece of mail matter in 1923, as estimated after an exhaustive study by the Post Office Department, was as follows: first class, 2.06 cents; second class, .7 cents; third class, 1.17 cents. The first class charge, even on this basis, is three times that of the second class.

The fiscal aspects of this postal rate structure are significant. The United States Post Office does not have a system of cost accounting such as is considered essential in any up-to-date private business. However the national interest in this problem is such that several special investigations have been made; the last, made by the Post Office Department and completed in 1924,[1] presents an analysis of revenues and expenditures for the year 1923. The following figures are selected as pertinent to our present inquiry.

[1] *Cost of Handling Mail Matter*, 68th Congress, 2d Session, Senate Document No. 162.

ALLOCATION OF REVENUES AND EXPENDITURES. UNITED STATES POST OFFICE, FISCAL YEAR 1923. (in millions)

Class of Service	*Revenues*	*Expenditures*	*Loss*	*Gain*
Paid first class	$272	$191	—	$80
Second class	31	106	$ 75	—
Third class	44	60	16	—
Fourth class	121	128	7	—
Other services	57	87	34	5
Total	$525	$572	$132	$85
Net loss*			$47	

* Certain unassignable and unrelated items are disregarded.

The first class business shows a profit of over 40 per cent of the cost, the result of charging 85 cents a pound or 2 cents a piece for a service which costs only 60 cents a pound or $1\frac{1}{2}$ cents a piece. On the other three classes there is a loss, moderate on the fourth class, considerable on the third, and very heavy on the second class. Here the cost is more than three times the revenue. The post office collects $2\frac{1}{4}$ cents per pound for a service which costs it $7\frac{7}{10}$ cents to perform.

The government has a monopoly of the business of carrying first class mail, under the protection of which it is able to charge high rates and make a large profit. But the profit is wiped out by the loss on the second class. The result is a subsidy to the periodical publishers at the expense first of the users of first class mail service and second of the general body of taxpayers, who have to make good the deficit on the postal business as a whole. This differentiation has many interesting aspects, political and social as well as economic, into which we cannot enter. For example, it means a heavy discrimination against books in favor of newspapers and magazines.

The postal deficit. Throughout the world in general the postal business is operated, not for a profit, but with the idea of making the revenues about cover the cost. In the United States the post office has always been operated at a loss, and the loss is actually considerably greater than the official figures show on account of the absence of capital charges in the accounts. The annual deficits must of course be met out of the general funds raised by taxation.

It is not likely that the second, third, and fourth classes could ever be made self-supporting, whatever the rates, so long as they are in competition with private enterprise. But it would be possible so to adjust the rates in each class as to produce for all services except the first class a moderate deficit only, which could be balanced by the profit on the first class.

The case for government ownership. There are those who believe that a considerable extension of the field of government industry would be of economic advantage to the people. This sentiment has appeared in connection with various topics which we have already studied. It is especially likely to present itself as the remedy for abuses which are found to attach to industrial combinations and monopolies and as a solution of the problems of the railroads and other public service enterprises. The argument in favor of government ownership lays stress upon the discriminations and unfair practices which accompany private competitive enterprise, and it becomes still stronger when private monopoly arises. Attention also is called to the large sums paid for interest on private capital and the large profits made by private entrepreneurs, and it is claimed that these gains might as well accrue to the whole people.

Considerations against government ownership: interest and profits. It will be recognized at once that the argument for government industry rests at certain points upon misunderstanding of fundamental economic principles with which the reader is now familiar. Interest is the reward for the service of waiting, which is essential to the accumulation of the capital equipment required by modern methods of production. So long as saving is left to private initiative interest will have to be paid; the government cannot get capital without interest any more than a private entrepreneur. If attention be called to the fact that government bonds bear a lower interest rate than industrial investments, the answer is that the financial difficulties which are the heritage of the World War have already deprived many governments of this advantage, and the whole advantage would tend to disappear in proportion as government became involved in the risks of industry. If the taxing power is relied upon to make good the losses of unsuccessful government

enterprises and so to enable the government to borrow at lower rates, this is merely robbing Peter to pay Paul.

The same general considerations apply to the alleged saving of profits. Profits, as we know, are paid for the service of risk bearing, and we have concluded that the payment, considering losses and gains together, is distinctly moderate. If the risks of industry were transferred from the shoulders of private entrepreneurs to the government, they would still have to be paid for. Government cannot take the profits and avoid the losses. The government industry would have to charge enough to enable it to make experiments, assume risks, and stand the inevitable losses. There is no reason to believe that the cost of risk bearing would be any less if undertaken by the government, and it might well be more, unless indeed risks were avoided and industry allowed to stagnate, which would certainly be no saving in the long run.

Discrimination. Again, while a certain degree of discrimination in prices and otherwise is an accompaniment of private competitive industry and still more of those industries which tend toward monopoly, it is an over-optimistic position to assume that all such discrimination would disappear under public ownership. Few private businesses make such glaring discriminations as, for example, the United States Post Office. First class patrons are, under the protection of the government's monopoly, charged much more than their service costs. To the periodical publishers is given a valuable service at a small fraction of its cost. Residents of remote thinly settled districts are given favors at the expense of those who live in other sections. All of this may or may not be justified, but at any rate great discrimination is present.

Politics. There are moreover certain positive weaknesses inherent in government management of industry. There can be no doubt of the fact that government management is generally less efficient than private business. In the first place there is weakness in personnel and discipline. Politics controls, more or less, the choice of employees, both managers and subordinates. There is no disguising the fact that public employees are commonly chosen on grounds of political affiliation or service rather than because of

fitness for the work. Examples are a commonplace in America in all grades of government, and the practice is accepted almost as a matter of course. This means necessarily that the government has less able managers and a less efficient labor force. Politics also interferes with discipline. Those in charge hesitate to enforce adequate control over their subordinates or to discharge the incompetent, for fear of political protest. For the same reason the managers have often been prevented from installing approved labor-saving machinery, cost accounting systems, or up-to-date business methods. Thus is efficiency sacrificed and industrial progress hindered.

Political sentiment also tends to prevent the payment of salaries sufficient to attract into the public service the most capable managers and technical experts. One of the serious handicaps with which the head of any government service has to contend is the constant drain of his most competent subordinates into private business, attracted by the lure of higher pay.

The motive in business. Public industry is weak as compared with private, with respect to the motive of the managing head. The private business man puts the utmost of his skill and energy into his business because he has a personal stake in it. He has risked his own capital; he faces the opportunity of large profits and also the danger of loss; his economic well-being and the comfort of those dependent on him hang upon the success of his enterprise. Even if he be only the hired manager of a corporate business, he finds the same standard held up to him by the stockholders, who look to him to protect their capital and produce profits for them. His position and his fortune depend on making a success of the business under his charge. Self interest is the spur which drives the private business man to make his business efficient.

To the public official this motive is weak or absent. He has not risked his own capital in the government industry. If profits are made they will not accrue to him, and he will not have to bear the losses of failure. The continuance of his position and his further advancement do not depend upon the judgment of a critical body of stockholders upon the efficiency with which he has managed the

business. His position is dependent upon pleasing the general body of voters or the leaders of his political party. This approval is to be won by easy-going treatment of his subordinates and showy service to the public, rather than by the hard test of business efficiency.

General conclusion. We have not the space in this book for a more complete analysis of the economic weaknesses of government industry. Enough has been said however to justify the general conclusion against government industry on economic grounds. The fact is that government industry is almost never as efficient as private. Few public industries make any profits; generally they are run at a loss, which has to be made up by taxation. In the somewhat rare cases in which a government undertakes to make a profit from an industrial enterprise, and succeeds, it will usually be found in possession of a monopoly; otherwise it would be driven out of business or forced to operate at a loss by the competition of the more efficient private enterprise. This is the case with the French tobacco industry, with most government salt industries, and with the United States first class postal business. Government simply cannot produce commodities or services at so low cost as private industry, for reasons which are now apparent. The public interest demands that goods be produced at the lowest possible cost. Nothing is gained by wiping out private profits if the result is a smaller volume of goods at higher cost. Except when an industry is required as an incident to the government's primary functions, the general conclusion is therefore clearly against government ownership.

The case of monopoly. But while this is the general conclusion, there is one exception — an exception of great importance and one which reopens the controversy in the only field in which it is likely to play any significant practical rôle in America. The question of government ownership in the ordinary fields of industry in which competitive private business now functions is not a practical one in this country. But there are certain lines of industry, as we have learned, in which there is a strong natural tendency toward monopoly. We have also reached the conclusion that private

monopoly is not generally in the interest of the public. When monopoly appears to be economical or inevitable it is therefore a fair question whether the evils of private monopoly may not outweigh the weaknesses of government ownership. This question arises particularly in connection with the great public utility enterprises, the railroads, the telegraph, the telephone, the express business, and such local public utilities as street railways, water supply, gas, electricity, etc.[1] It is with regard to these public utility industries that the question of government industry is a live one in the United States.

Conclusion as to public utilities. The question of government ownership of public utilities cannot be settled off hand by mere reference to general principles. One conclusion to which our previous study[2] has brought us is that the public will not tolerate unregulated private monopoly in the production of any vital necessity. The choice is between private industry with government regulation and government industry. On this question we are, in the face of the general principles which have been established, justified in placing the burden of proof upon the advocate of government industry. Our attempts at government regulation have thus far been generally timid and bungling. We are now it would appear upon the threshold of our first real test of this program. If it should fail, the alternative of government ownership will still be open to us. If we should first adopt government industry and later learn to regret our choice, the return to private industry would be extremely difficult.

This question will probably be settled gradually and piecemeal, by experiments here and there. We have now publicly owned and privately owned municipal water systems. Eventually experience in the different cities should enable us to judge which plan has the greater advantage to offer the public. Conclusion may likewise come in the field of street railways, gas and electric supply, and other municipal utilities. During the World War we had a test

[1] The nature of the public utility, with special reference to the railroads, is discussed in Chapter XXIX.

[2] See Chapter XXX.

— though admittedly far from perfect — of government management of the railroads. The lessons to be drawn from this experience are not entirely clear, but much of pertinence was learned, and it is safe to say that public sentiment in the United States is less inclined to favor government ownership of the railroads than before this experience. Similar results followed the government operation of the telegraph and telephone systems during the war. Public sentiment for government ownership — in America at least — has suffered a considerable setback within the past generation. The prevailing attitude today inclines to further careful test of the program of government regulation of those industries which show a monopolistic tendency.

Suggestions for Further Reading on Part VI. The Economics of Government

I. Public Finance. General Texts

Lutz, H. L. *Public Finance.* New York, 1924
Jensen, J. P. *Problems of Public Finance.* New York, 1924
Shirras, G. F. *The Science of Public Finance.* London, 1924
Robinson, M. E. *Public Finance.* New York, 1922
Bastable, C. F. *Public Finance.* Third edition. London, 1903
Plehn, C. C. *Introduction to Public Finance.* New York, 1920
Adams, H. C. *The Science of Finance.* New York, 1898
Seligman, E. R. A. *Studies in Public Finance.* New York, 1925
Bullock, C. J. (Editor) *Selected Readings in Public Finance.* Third edition. Boston, 1924
Dewey, D. R. *Financial History of the United States.* Eighth edition. New York, 1922 (See Table of Contents)
Noyes, A. D. *Forty Years of American Finance.* New York, 1909

II. Taxation

Wells, D. A. *Theory and Practice of Taxation.* New York, 1911
Armitage-Smith, G. *Principles and Methods of Taxation.* New York, 1906
Lyon, W. H. *Principles of Taxation.* Boston, 1914
Brown, H. G. *The Economics of Taxation.* New York, 1924
Seligman, E. R. A. *Essays in Taxation.* Ninth edition. New York, 1921
Ely, R. T. *Taxation in American States and Cities.* New York, 1888

HOWE, F. C. *Taxation and Taxes in the United States under the Internal Revenue System, 1791–1895.* New York, 1896
SMITH, H. E. *The United States Federal Internal Tax History from 1861 to 1871.* Boston, 1914
LUTZ, H. L. *The State Tax Commission.* Cambridge, 1918
SELIGMAN, E. R. A. *The Income Tax.* New York, 1911
HAIG, R. M. (Editor) *The Federal Income Tax.* New York, 1921
WEST, M. *The Inheritance Tax.* New York, 1908
FILLEBROWN, C. B. *The A B C of Taxation.* New York, 1909
POST, F. L. *The Taxation of Land Values.* Fifth edition. Indianapolis, 1915
National Tax Association. *Proceedings of the National Conferences — Annual from 1907*
Bulletin, Monthly from February, 1916
National Industrial Conference Board. *Taxation and National Income.* New York, 1922
Tax Burdens and Exemptions. New York, 1923
Tax Burdens and Public Expenditures. New York, 1925.

III. CUSTOMS TARIFF. PROTECTION

TAUSSIG, F. W. *Tariff History of the United States.* Seventh edition. New York, 1923
PAGE, T. W. *Making the Tariff of the United States.* New York, 1924
TAUSSIG, F. W. *Some Aspects of the Tariff Question.* Cambridge, 1915
ASHLEY, P. *Modern Tariff History — Germany, United States, France.* New York, 1911
ASHLEY, W. J. *The Tariff Problem.* London, 1911
PATTEN, S. N. *The Economic Basis of Protection.* Philadelphia, 1890
DIXWELL, G. B. *The Premises of Free Trade Examined.* Cambridge, Eng., 1881
JORDAN, D. S. *The Fate of Iciodorum.* New York, 1909
PIERCE, F. *The Tariff and the Trusts.* New York, 1913
STANWOOD, E. *American Tariff Controversies in the Nineteenth Century.* New York, 1903
PIGOU, A. C. *Protective and Preferential Import Duties.* London, 1906
DAWSON, W. H. *Protection in Germany.* London, 1904
MEREDITH, H. O. *Protection in France.* London, 1904

IV. BUDGET, GOVERNMENT BORROWING, WAR FINANCE, ETC.

STOURM, R. *The Budget* (English translation). New York, 1917
CLEVELAND, F. A. *The Budget and Responsible Government.* New York, 1920
BUCK, A. E. *Budget-Making.* New York, 1921
PLEHN, C. C. *Government Finance in the United States.* Chicago, 1915

DAWES, C. G. *The First Year of the Budget of the United States.* New York, 1923
CLEVELAND, F. A. *Municipal Administration and Accounting.* New York, 1919
GRICE, J. W. *National and Local Finance.* London, 1910
ADAMS, H. C. *Public Debts.* New York, 1887
HOLLANDER, J. H. *War Borrowing.* New York, 1919
NOYES, A. D. *Financial Chapters of the War.* New York, 1916
BOGART, E. L. *Direct and Indirect Costs of the Great World War.* New York, 1919
BASS, J. F. and MOULTON, H. G. *America and the Balance Sheet of Europe.* New York, 1921
MOULTON, H. G. *The Reparations Plan.* New York, 1924
World Peace Foundation. *Pamphlets on Reparations.* Boston
THOMPSON, C. D. *Public Ownership.* New York, 1925
MEYER, H. R. *Municipal Ownership in Great Britain.* New York, 1906

PART VII

PROBLEMS OF LABOR

CHAPTER LI

UNEMPLOYMENT

Origin of the labor problem. Before the rise of machine industry the worker's tools were usually in his home. He was in some senses an entrepreneur, setting himself to work upon a product which brought him into direct contact with the market. It was customary also for the worker to hold a plot of ground, which supplied him with a supplementary source of income and which could serve as a resource held in reserve for slack periods of trade. This position has been revolutionized by the development of the factory system. In the early stages of the Industrial Revolution, the capital instruments employed in the typical enterprise became so costly that it no longer lay in the power of the average workman to own them. The use of natural power necessitated the concentration of these instruments in factories, thus destroying the utility of the laborer's home to serve as his workshop.

As a result of these changes, the average wage-earner of today owns practically no capital. There remain a few of the skilled trades in which the laborer as a customary thing supplies himself with capital instruments, but these are exceptions to the general rule. In many of the older countries — in England for example — this process was accompanied by another which worked in the same direction; namely, the separation of the laborer from the land. For reasons which we need not discuss again, an agrarian revolution, approximately paralleling the series of inventions which rendered obsolete the workman's simple tools, swept away the system of small land holdings which had characterized the earlier agrarian economy of England. In other countries, where this change was not so noticeable, and especially in the United States, where an abundance of land postponed its effects until our own time, the workingman of today is after all in little better case.

He is born landless and is rarely able to scrape together the money required to buy a portion of the earth's surface. And this is a condition which is bound to extend as the increasing pressure of population on the land areas of all industrial countries works out its effects in rising rents and rising land values. Moreover the development of that aspect of the Industrial Revolution which we call the factory system has brought it about that small land holdings scattered over the countryside are too remote from the scene of the laborers' daily work to serve as a source of supplementary income.

Extent of unemployment. The menace of unemployment is the greatest of the problems confronting the industrial wage-earner, since it strikes at the very foundation of his life. The magnitude of the problem is disclosed in the statistics compiled by trade unions, governmental bodies, and the various research organizations of the country. In Massachusetts, Pennsylvania, New York, and Illinois, it has been found that the average number of days lost by all classes of wage-earners is 42 per year, about 14 per cent of the working time. Among organized workers, consisting chiefly of skilled men, 10 per cent of the membership is unemployed in normal times and almost 20 per cent in years of depression. When we take into account also the unemployed among unskilled and unorganized workers, we arrive at about the same ratio of normal unemployment; *i.e.*, 15 per cent of all wage-earners, or 2,000,000 workers, continuously idle in normal times.

But the most serious phase of the problem of unemployment appears in connection with that feature of the modern industrial system which we call the business cycle. During the crisis which comes at the peak of the price level and throughout the business depression which follows, unemployment spreads over the entire range of industrial enterprise, affecting workers in the most stable industries. Thus it is estimated that two and a half million workers were suddenly thrown out of employment in the United States during the crisis in 1915 and that as many more were put on part time. At least three and a half millions were totally unemployed following the crisis of 1920.

The unemployed. The unemployed, as the word is used here, consist of workmen unable to find employment though seeking it and quite competent to perform work. It does not include men on strike or those who are unemployable. With the unemployable we shall not concern ourselves in this study further than to say that their condition is due to some weakness in themselves, some moral or physical flaw which renders them useless as productive instruments in society. Their presence constitutes a special social problem, which calls for its own specific remedies: segregation of these moral and physical derelicts from the rest of the population, rehabilitation where feasible, permanent segregation in the hopeless cases. We must be on our guard however against concluding that the army of the unemployed is made up principally or to any significant extent of such people as these. At any given time only a small fraction of the men out of work is unemployable. Strikes also are but an insignificant cause of unemployment. Most of the unemployed are both able and willing to serve society as producers. But they are prevented from applying their labor in this service through the operation of forces extraneous to themselves and beyond their control. It is with this group that the economist is concerned when he treats of the problem of unemployment.

The causes of involuntary unemployment. Inherent in the unstable and dynamic character of our industrial system are the causes of involuntary unemployment. They are due to recurring unforeseen changes in the productive system, which upset the equilibrium between the supply of labor and the demand for labor in different occupational groups and industries. We may throw these causes into three general groups: (1) those associated with the progress of invention, (2) those due to seasonal fluctuation of demand, (3) those involved in the business cycle.

Invention. Improvements in the technical processes of industry are inspired by the constant pressure to reduce the cost of production. In the common run of cases, these savings of cost take the form of reduction in the wages cost of producing a given quantity of goods. This occurs either through a reduction of the

number of men employed in producing this quantity of goods or through the substitution of a cheaper for a higher grade of labor. The same effect may be seen more clearly if we view the matter from the standpoint of society as a whole, for it is obvious that no invention is a real improvement in the social sense unless it economizes the labor expenditure involved in a productive process. The immediate effect of invention therefore is to reduce the demand for the labor of some particular group of workers. It is true that this effect in the average case is neither violent nor revolutionary. The progress of invention has become for the most part a matter of slow accretions rather than of sweeping changes. The new improvement usually affects only a relatively small group of workers, and the whole process has become so gradual that the laborers have the opportunity to adjust themselves to the changed conditions without much loss. It is in those cases, to be found in the record of the Industrial Revolution, when society has adopted a thoroughgoing change in technical progress that the immediate effect of invention upon the demand for labor is made most apparent. Of such significance was the invention of the power loom, the nail-making and bottle-blowing machine, the linotype machine. Though the ordinary changes in machine equipment are not so revolutionary as these, they differ in their immediate effect upon labor only in degree. In all cases the new machine alters to some extent the conditions of demand for labor, and the continuous and unceasing progress of machine equipment, though it operates in one direction to temper the effect of invention upon the worker, injects an element of insecurity and uncertainty into the lives of all industrial wage-earners.

In general the working classes of the world have taken this view of the significance of invention in their lives. They look upon the new machine with suspicion and tend either to oppose its introduction, where they have the power to do so, or to cripple its functioning by operating it in a spirit of resentment and antagonism. This behavior of labor is so opposed to the long-run interests of all members of society that many efforts have been made to convince the workers that their attitude toward invention is

based upon false reasoning. It is pointed out that the ultimate effect of invention must be to lower the price of goods in the market; a greater quantity of goods will be sold at the reduced price, and the labor set free by the cost-saving device will be reabsorbed through the expansion of industry. If this is true no permanent unemployment need occur, and there will be a positive gain for labor, together with all other members of society, in the increasing plenty of want-satisfying goods.

Concerning the beneficial effects of progress in the arts of industry upon society as a whole and, in the long run, upon the working classes, as thus pointed out, there can be no argument. The reader will recognize this as a consequence of the law of proportionality, which was investigated in Chapter XXXIX. But we are considering here the effect of invention upon a particular group of workers. From this point of view there is no assurance that the workmen immediately affected by a new machine will recover what they lose through the reduction in the demand for their labor. Sales of goods expand as their prices fall, but they do not inevitably expand sufficiently to absorb in the same industry all the labor thrown out of work by the invention. This depends upon the elasticity of the purchase of the good in question. Even if sales increase so much that the industry provides work for as many men as before, these need not be members of the same labor group which was employed before the change in technical process. Unskilled or semi-skilled workmen may take the place of a higher grade of labor or, to state the matter otherwise, the demand for the labor of one group may be reduced and that for the labor of a different group increased by the invention. Moreover the workers who have been adversely affected can profit by the decline of market price only to the extent that they consume the product of their own industry, and this, as we have seen, is not the rule in an economic system characterized by refined division of labor. If the invention works considerable change in the processes of industry, it is probable that some of the wage-earners must seek employment in other lines of activity, either in the same locality or elsewhere, or must drop into a lower labor group. They suffer

unemployment while they are adjusting themselves to the altered conditions and may recover their position in the world of industry only at a permanent sacrifice of real income. On the whole however we must conclude that the progress of industry is today of no great importance as a cause of unemployment.

Seasonal fluctuations. These are so diverse in nature and conform so closely to the peculiarities of the individual enterprise which they affect that they must be studied in each case as technical phases of particular industries. In general however their effects appear in the form of an unequal distribution throughout the year either of the demand for the product of the enterprise or of the supply of raw material which it utilizes. For example in the clothing trade there is a sharp concentration of demand for certain types of clothing in brief periods of the year. The demand for the product of coal mines varies markedly with the seasons. The erection of buildings is seasonally affected both by changes in demand and by climatic conditions which limit operations. The manufacture of foods — such as canned fruits and vegetables — fluctuates in sympathy with the seasonal nature of agriculture. These are some of the more visible effects of this factor, but its influence is widespread. Some enterprises are affected indirectly by reason of their position as subsidiaries to other industries whose market is disturbed by seasonality of demand; an example is the cement industry, which feels the effects of climatic control of building operations. Manufacturers of specialties, who sell a finished product to the ultimate consumer, find that the demand for their output fluctuates as a consequence of changes in fashion or in harmony with the forces of habit and custom. Trade union statistics for the State of New York covering the years 1904–1915 show an average unemployment in March 75 per cent higher than in September. In Massachusetts similar statistics show that the average in March was 80 per cent higher than in September during the years 1908–1919. It is estimated for the country as a whole that one and a half million more workers are unemployed in the worst month of the year than in the best month.

The labor reserve. It is the tendency of industries subject

to wide seasonal variations to attract to themselves a disproportionate number of workers as compared with the more stable enterprises. The attractiveness of the relatively high wage rate, amplified by bonuses, overtime pay, and other devices upon which employers rely to stimulate production during the season of insistent demand, causes the workers to lose sight of the obvious fact that their annual earnings must be reduced by periods of enforced idleness. The supply of labor permanently attached to such industries tends to be equated to the highest seasonal demand. In other words, there is a *reserve of labor* in these enterprises — an excess of workers over and above the number who could be continuously employed if the seasonal variations were removed. This labor reserve is drawn upon to equip the industries during the period of intense activity and is refilled by the discharge of workers as the volume of sales diminishes and business grows dull.

Examples of the labor reserve may be found in almost every industry which is sharply seasonal. Experts who have sought an explanation for labor unrest in the bituminous coal industry declare that two hundred thousand coal miners would be permanently unemployed if this industry were so stabilized as to make the average production continuous throughout the year. In the clothing trades large numbers of wage-earners would remain idle if the total annual output could be distributed uniformly over the seasons. In the sharply seasonal occupations of harvesting the staple food crop and gathering the produce of orchards and vineyards there are at least half a million more men employed than would be needed if these fluctuations could be removed.

The social injury. The social injury which flows from this state of affairs is the loss of labor power, as measured by the average idleness of the working force. Until the time comes when all human needs are satisfied and all productive opportunities exploited to the full, an excess of labor in any one line of activity can occur only at a sacrifice of product in other departments of the industrial structure. If the reserve of labor in a given industry is so large that one third of all who work must, on the average, be idle all the time, the loss of their work must be charged as one

of the costs of operating the industry, regardless of the fact that the employers bear a wage expense equivalent to only a part of this social loss. The fact that the unemployment is shared piecemeal by the entire labor group and not thrown entirely upon the extra third in no wise qualifies the principle involved, though it does conceal the loss and may act as an obstacle to its correction.

Labor's loss through seasonal fluctuations. The workers in the seasonal industries are not concerned with these considerations of broad social welfare but only with the effect of fluctuating employment upon themselves. This effect is two-fold; in the first place it makes the actual earnings of the worker less than would be indicated by the wage rate; secondly it diminishes the utility of these earnings. So clear is the first of these effects that it could be dismissed without mention except for the abundant evidence that it is imperfectly understood, both by the general public and by the wage-earners themselves. Whenever a strike occurs in a seasonal industry, such as coal mining or the building trades, the public at once seizes on the relatively high daily rate of wages as evidence that the workers are presenting exorbitant demands. It is also true that the incoming generation of workers, impressed by the higher daily rates in seasonal as compared with stable industries, neglects to discount the higher rates by translating them into normal yearly earnings. It is for this reason that seasonal industries tend to build up a labor reserve.

The second effect of seasonal unemployment is not so obvious as the one just discussed. Quite apart from its effect upon the amount of earnings, persistent recurrence of unemployment reduces the utility of the wage-earner's income. The reader will recall the law of diminishing marginal utility (Cf. Chapter XVI). To deprive the laborer of his income during certain periods, in order to add somewhat to an already ample stream of wages at another period, must reduce the utility of the income. Periodic unpredictable unemployment operates in this way. Wages are abstracted from a period when their marginal utility is high in order to increase wages in a period when their marginal utility is relatively low.

Effects of the business cycle. The different phases of the business cycle have diverse effects on the welfare of labor. During the prosperity period there is a tendency for wages to rise less rapidly than other prices, with the result that the real income of the wage-earners falls. Nevertheless it is probably true that the net effect upon the workers of this phase of the cycle is good. This is a period when unemployment is reduced to a minimum. Members of the family not usually employed are drafted into service. Overtime pay at rates higher than the ruling scale of wages, bonuses, and other perquisites are bestowed upon the worker by the employer in his effort to maintain a stable working force.

The problem of unemployment appears at the time of the crisis. This condition is found not only in one enterprise or group of enterprises but in all types of industries throughout the entire country. Consequently the unemployed can find no escape from the disaster in other industries or other places. They have no alternative but to remain idle and attempt to tide their families over the "hard times" by drawing upon the slender reserves accumulated during the period of prosperity. The savings of years are sacrificed, household effects are surrendered to the pawnbroker, homes are mortgaged. In the more serious cases large numbers are forced into pauperism. Charitable organizations in the great cities mobilize their resources to keep groups of industrious and effective working people from starving. The dockets of the police courts show that many of the sufferers are driven to petty crime. It is a time of unrest, when radical ideas take root and flourish among the industrial workers.

During this period, as we have learned, the industrial structure is undergoing readjustment. In time it reaches a new point of equilibrium and business activity begins gradually to quicken. Whether the distress of the working population will be long drawn out or of brief duration depends upon the extent to which the various lines of industrial activity had been drawn out of adjustment to each other during the period of rising prices. If important groups of industries have been much overdeveloped in proportion

to other groups, a long and painful period of readjustment will be required to reëstablish the balance. During this period the workers must shift in considerable numbers from certain industries to others and from locality to locality, until they are redistributed upon a new basis over the field of industrial enterprise. The crisis of 1915 was caused by rather artificial forces resulting from the outbreak of war. The industrial structure of the country was fundamentally sound, and, although two and a half millions of workers were thrown into unemployment by the crisis, their period of idleness was of brief duration. The crisis of 1920 was much more serious, both in respect of its initial effects and with regard to the duration of unemployment. At least three and a half million laborers were thrown out of work by the crisis, and many of these were continuously idle for over a year. In cases where the business depression which follows the crisis is severe and protracted, the workers suffer other losses than the mere cessation of income. The revival of business activity will not restore the status which existed before the crisis. Some groups of the working population must break the ties which bind them to the communities in which they have been living and must bear the costs of establishing themselves and their families in a new environment. Others will be required to drop out of the particular crafts or trades in which they have acquired skill and to undertake a new function, often on a lower basis of income.

Defensive policies of labor: the full crew. The forces which cause unemployment in our dynamic world are beyond the control of the workers who feel their adverse effects. However it is but natural that they should seek to protect themselves, using for this purpose any weapons which lie in their power. Of this nature is the insistence by the trade union upon what is called a "full crew." Agreements between trade unions and employers in the building trades often specify that each group of craftsmen must be accompanied by a minimum number of apprentices or helpers. The labor organizations on the railroads for a long time demanded that each train be staffed by a certain number of workers of various classes without much regard to the real utility of the

full crew in the circumstances of the case. The seaman's union advanced similar demands with regard to ocean-going vessels. The policy of these two groups of transport workers has now found its way into the law of the land and is enforced upon employers by the government. Other examples of this line of trade union policy could be given. It is not contended that the demand for the full crew is based solely upon the desire to create employment; in many circumstances it merits consideration on grounds of sound public policy. Nevertheless it is the general tendency of organized labor groups, when they have the power, to insist that each enterprise employ just a few more men than are really needed in the interests of economic efficiency. Such demands are in harmony with the wage-earner's habit of looking upon economic enterprises as job-creating agencies.

Labor's proprietary interest in the job. Fear of unemployment will explain other practices of the trade unions, which are almost incomprehensible to those who are not members of the wage-earning group. The importance of the job to the worker is so great that he develops a proprietary attitude toward it whenever conditions are favorable. Strongly organized groups which have entered into collective agreements with their employers take the view that the jobs are owned jointly by employer and worker. Neither party to the labor contract has the right to dispose of these jobs without consulting the other party. An understanding of this view of the situation throws light upon the behavior of men when on strike. As viewed by the law of our country and by that part of the general public which shares the employers' philosophy of property rights, a strike is a voluntary surrender of employment by a group of workers acting collectively. This action has freed the employer from any obligation toward his former labor force and has authorized him to offer employment to anyone in search of work. But the strikers do not acquiesce in this opinion. They have quit work with the intent of retaining the job; the strike is merely a contest over the terms of employment between the two parties who jointly control employment in the enterprise. Until this contest is settled, the enterpriser does not have a free

hand in filling the vacant places. Any worker who accepts employment during the strike appears in the light of a job stealer, a "scab"; and against him aggressive measures may be employed with propriety. Here we have a clash of ideas so fundamental that it is frequently impossible to reconcile the two points of view without use of force. Hence labor disputes often result in violent conflict between the organized workers on the one hand and the employer or the non-union men on the other.

The "lump of labor" theory. But the most far-reaching effect of the fear of unemployment upon the conduct of the workers is seen in the general disposition of labor — the unorganized as well as the members of trade unions — to oppose measures designed to increase production. It is impossible to calculate how much society loses through the make-work and time-wasting devices which laborers so frequently adopt. The make-work idea, with reference to the problem of unemployment, takes the form of the "lump of labor" theory, which asserts that there is at any given time only a certain amount of work to be done. The problem of each laborer is to get his share of this quantum of employment. If one man through extra exertion does the work of two ordinary workers, he drives someone else into unemployment. If all workers discharge functions with unusual efficiency, they will all be forced into idleness a part of the time. Employment is a thing to be husbanded, economized, made to last. As the negro stevedore stated the matter, "What is the use rushing through work to get the chance of looking for more work?"

The lump of labor theory is fallacious. The amount of work to be done in the world is not fixed in quantity, but is indefinitely extensible. The different groups of workers create employment for each other, and a general increase in efficiency will merely lead to an increased plenty of goods and a higher average of income for everyone. On the other hand, a general indulgence in "ca' canny" and make-work can lead only to general impoverishment.

It is well to point out however that this disproof of the lump of labor theory does not demonstrate that make-work policies are always unwise from the standpoint of the particular men who

practice them. Though the wants of mankind considered as a whole are indefinitely extensible, each particular want is satiable. Where they have the power to control the supply of their type of labor, a group of workers may increase the regularity of their employment by decreasing their rate of output. On the other hand, it is often true that an increase in efficiency on the part of a group of workers will either force some of them into other lines of employment or reduce their wage rate. As long as no collective attempt is made by society to solve the problem of unemployment or to lighten its effects upon the wage-earners, it is not to be expected that any specific labor group will refrain from adopting in their own protection measures which react injuriously on other social classes and other groups of workers.

Remedies for unemployment. There are two ways of attacking the problem of unemployment. One is to assume that the abrupt changes in the demand for labor which cause the trouble arise from irremovable causes and to concentrate upon the problem of adjusting the supply of labor to the altered conditions of demand with as little delay and friction as possible. The other is to attempt to control demand by stabilizing industry.

Adjustment of the labor supply. Considering the first of these methods, it is to be noted that labor is not the only good subject to these fluctuations of demand. Yet in the case of other goods the adjustments of supply take place almost automatically and without much delay. When consumers' demand falls, the result is not a destruction of a part of the stock of goods on hand similar to the destruction of labor power which occurs when men are thrown out of work under similar conditions. Why do not similar adjustments take place with similar ease in the labor market?

Peculiarities of labor supply : impossibility of storage. One minor reason for this difference is to be found in the fact that the device of *storage* cannot be applied to labor. Most products of industry can be held in reserve for a better market during a period of temporary depression. They may deteriorate while in storage, but at all events they need not be destroyed utterly because the demand has fallen. But labor power is a peculiar thing in that

it cannot be stored at all. If the workman fails to sell today's labor today, he can never sell it; it has been lost beyond recall.

Rigidity of the wage rate. But there are two other characteristics of the labor market which tend even more powerfully to prevent a smooth adjustment of the supply to changes of demand. The first of these is the relative rigidity of the wage rate as compared with other prices; the second, the immobility of labor. When an industry is affected by some market change so that its demand for labor shrinks, the entire force may still find employment provided the wage rate is sufficiently flexible to fall at once to a point which will bring about a new equilibrium between supply and demand. A relative surplus of labor in one industry or in one locality as compared with another need cause no unemployment if the wage rates are sufficiently pliant. Even the unemployment caused by seasonal and cyclical fluctuations could be much reduced, if not entirely avoided, if the wage rate were to move quickly into adjustment with the new conditions. The wage rate however is affected by many factors which do not influence the more impersonal markets for other goods. Emotional influences play their part in preventing sudden or drastic changes in the rates; as do also the unconscious forces of habit and tradition. The result is a price which moves sluggishly and haltingly in response to market changes. It follows from what has been said that any device which increases the flexibility of the wage rate will act as a corrective of unemployment, and *vice versa*.

Immobility of labor. The other factor which prevents an automatic adjustment of supply to demand in the labor market is the immobility of the working population. This appears in two different forms, *functional* and *geographic immobility*. There are times, as we have seen, when the demand for labor of a particular kind changes in proportion to the demand for another kind; such, for example, is the effect of invention on the labor market. At other times forces are at work to dislocate the geographical distribution of industry and to change the relative demands for labor in different localities. Adjustment to the former of these changes requires that some workers abandon the particular productive

function which they are discharging and seek out new lines of activity. In the second case a portion of the working population must move from one district in the country to another, in order to produce a more perfect geographical distribution of the labor supply. But neither of these adjustments can be worked out in the labor market without great friction. When the workers have devoted their early years to acquiring skill in a particular craft, they tend to cling to this line of activity regardless of changes in the demand for it; to do otherwise involves for them a sacrifice of earning power which they are loath to endure. On the other hand, when laborers have become settled in certain communities, to move elsewhere not only requires a severing of the ties of friendship and neighborliness which always oppose such changes, but involves expense of travel which wage-earners are often in no position to bear. Furthermore it is frequently true that the laborers in one locality are ignorant of employment conditions in another and possess no sure and inexpensive means of obtaining this information. For these reasons, among others, the members of the labor force tend to become so fixed in their various functional lines and in their different localities that redistribution is retarded. We conclude then that anything which promotes the *versatility* of the average wage-earner will act as a corrective of unemployment and that any device which increases his *mobility*, in the geographical sense, will have the same effect.

Labor exchanges. Unemployment may be somewhat reduced by agencies which act as clearing houses for information regarding the demand and supply of labor. This is the function of the labor exchange. A system of labor exchanges maintains an office in each industrial centre, which registers and classifies the idle workmen and the available jobs in the locality with the object of distributing the unemployed labor over the entire area with as little delay as possible. If the government does not undertake this function, it will be discharged, as a profit-making enterprise, by private initiative.

Private labor exchanges. The private exchange however has been found to develop into an evil unless controlled. The man-

ager has no interest in reducing the amount of unemployment; his income is drawn from fees charged the idle workmen who register at his agency. Instances have been known of collusion between the owners of labor exchanges and the foremen or employment managers of industries. The foremen agree to discharge all newly employed workmen as speedily as possible so that a steady stream of applicants may pass through the exchange on the way to the job. Many times the fees charged for the service of the exchange have been exorbitant, especially during periods of acute distress. Private exchanges have also sent workmen on futile journeys to distant cities equipped with cards of introduction to fictitious employers. They have worked in collusion with loan sharks who advanced money required for railroad fares at usurious rates on chattel mortgages. Experience with this institution has proved that it is unsafe to entrust the solution of a social maladjustment to a private enterpriser who draws his profit from its continuance.

Public labor exchanges. Thirty-eight American states have statutes regulating the private labor exchange with more or less rigor. But the only effective solution of the problem is for the government itself to assume the function of supplying idle workmen with information regarding opportunities for employment. Great Britain has a system of publicly owned and managed employment exchanges covering the entire country and giving its services free of charge to the workmen. This institution not only serves to distribute the labor supply during periods of unemployment but is a valuable agency for the collection of employment statistics to be used as the basis for a more thoroughgoing solution of the problem. About twenty-three states in our country maintain public employment exchanges. In the majority of cases however these are virtually impotent to discharge their function because of insufficient financial provision. Efficient systems of labor exchanges are to be found in the great industrial States of Massachusetts, Ohio, New York, Pennsylvania, Wisconsin, and Illinois. As a part of its war policy, the United States Government set up a nation-wide system in 1918. Between January, 1918, and

March, 1919, 5,323,509 unemployed workmen were registered, while applications from employers totalled 10,164,000. Of the registered workmen almost 5,000,000 were referred to positions, and of these three and three quarters millions were successful in obtaining employment. The wave of economy which swept the country after the war destroyed this efficient system by cutting off its financial provision. The experience of the world with labor exchanges is conclusive proof of their effectiveness and their economy. Their saving of lost time and vacant jobs, to say nothing of their reduction of the burden of unemployment on the working classes, is sufficient to justify any expenditure that society may be obliged to make in maintaining them.

Stabilizing the demand for labor: public works. As already stated, a different type of solution for the problem of unemployment consists in attempts to stabilize the demand for labor. It has long been urged that the government could accomplish something in this direction by managing its own employment of labor in such a way as to increase demand during normal slack periods in private enterprise. The routine construction work of local and central governments could be concentrated on periods of seasonal unemployment. Larger undertakings, such as reforestation, irrigation projects, or ship-building, could be carried through during the depression which follows a business crisis. It must be admitted that much of this pressure for public expenditure on the employment of idle workmen has in mind projects which are sheer waste, or worse. The government can acquire purchasing power only by reducing that of its citizens through taxation or borrowing. Waste of public money on projects which serve no useful purpose will not create additional employment but merely substitute objectless labor for that which would have enriched society.

The scheme of public works however may be of some benefit as a means of mitigating the effects of unemployment if applied in conjunction with more fundamental measures. If carefully planned and safeguarded against misguided philanthropy, a program of governmental expenditure on projects which are worth

while for their own sake could be pushed ahead in lean years and in months when private employment is slack. It is necessary for the success of such a program that it be based on reliable statistics showing the trend of seasonal and cyclical unemployment. On the basis of this information public work may be planned over a period of several years in advance, the necessary expenditure being financed out of reserves accumulated for this purpose from the ordinary revenues of government. The work should be done in the ordinary way, the workers being employed at the standard wage and under usual working conditions and hired on the basis of efficiency, not merely because they happen to be unemployed. It must be borne in mind however that many governmental undertakings are subject to the same climatic conditions which are the basic cause of seasonal fluctuations in private industry. Moreover any such program always involves the risk of the political chicanery and extravagence which causes us to distrust all forms of governmental activity in economic matters. On the general question of government industry, the reader should consult Chapter L.

Stabilization by private enterprise. Within recent years managers of enterprises in seasonal industries have taken steps to stabilize their demand for labor on their own initiative. The motive behind these attempts to reduce unemployment has been a desire to increase efficiency of labor and reduce overhead costs. Modern employers have learned that the constant fear of unemployment is a potent cause of the tendency of wage-earners to restrict output. Furthermore the turnover of labor in a given enterprise involves expense of hiring and training the workers which could be reduced by maintaining the labor forces at a constant level and retaining the same workers on the payroll for a considerable period of time. In the clothing markets of Chicago and Cleveland it has been found that the wage-earners are quite willing to guarantee a standard rate of production, provided the employers in turn guarantee continuous, or nearly continuous, employment. The experience of many large concerns has also proved that wage-earners will attach themselves permanently to that employer who can promise a lifetime job.

A detailed study of the devices by which seasonal industries are stabilized lies in the field of business management and outside the limits of our study. We shall not attempt to do more than summarize the more important of these devices in this place. (1) An industry whose major product is subject to fluctuating consumers' demand may take on *compensating side lines* whose market varies inversely with that of the major product. An example is the manufacture of felt hats or cloth caps by a concern specializing in straw hats, or the production of some staple product by a concern devoted primarily to the manufacture of Christmas specialties. (2) Somewhat similar is the development of *compensating markets*. Cement manufacturers have regularized their output by exploiting the South American markets where the seasons are the reverse of our own. Much can be done on a smaller scale along this line within the continental area of the United States. (3) Goods may be *made to stock* during the slack months. This device involves expense of storage and can be used with safety only in connection with types of product which do not deteriorate in storage and are not subject to capricious changes of fashion. (4) Sales methods may be adjusted to encourage *advance orders* with a long delivery time. By offering particularly favorable terms, a large concern specializing on the Christmas trade has succeeded in inducing many of its customers to place their orders more than a year in advance, thus allowing a stabilizing of production. Similar results have been obtained by many of the larger clothing manufacturers.

Unemployment insurance. There is as yet no hope that devices for increasing the mobility of labor and for regularizing the demand for labor will prevent unemployment entirely. We know of no practicable method of stabilizing the business cycle. Changes in technical processes, in the geographical location of industries, and in fashion will continue to upset the smooth functioning of the labor market as long as industry is controlled by private initiative. In highly competitive fields of production the adoption of many devices which would reduce the amount of seasonal unemployment is prevented by the lack of coöperation among

the different producers. Assuming then that a certain amount of unemployment is a normal and inevitable feature of our system of industry, is it possible for society to take upon itself the burden of loss which results from this evil?

The case for insurance. At first sight unemployment seems to conform to the tests of an insurable risk. It applies to a sufficiently large group of people and over a sufficiently diversified range of conditions to permit of the operation of the law of averages. By statistical study the normal rate of unemployment for each of the larger groups of industries can be determined with sufficient accuracy to make it possible to predict the number of unemployed and to provide in advance for their relief. The economic loss involved in unemployment is definite and ascertainable; it can be wholly or partially compensated by indemnifying the idle workman on the basis of his former earnings. Not only do these features of the unemployment hazard appear to make it amenable to the principle of insurance as a practical matter of business, but weighty arguments of a general social nature can be urged in favor of the policy. On grounds of justice it appears unfair that those least capable of bearing the burden and least responsible for its causes should be called upon to suffer. Furthermore the only real improvement in the situation consists in preventing unemployment or reducing its amount, not in expedients for distributing the loss occasioned by it. Little progress will be made in this direction as long as the loss is concealed from view; unemployment must be made a clear and definite expense to those who control industry and to the general body of society. Finally recognition of the fact that the risk of unemployment is a cause of far-reaching productive inefficiency among the workers argues that society would gain more in increased productivity from a policy guaranteeing the security of labor than it would lose through the expense of administering the policy.

The Ghent system. A tentative approach toward mutual unemployment insurance has been made by some of the stronger trade unions. Many local branches of the great national organizations provide for the accumulation of funds from which to pay

small pensions to members out of work. These systems are carefully safeguarded against fraud by tests of the authenticity of the worker's claim for relief and by setting a maximum limit to the duration of relief in individual cases. Since the amounts paid are small and not guaranteed by adequate reserves and since the system applies only to a small minority of organized workers and not at all to the unorganized, this plan has but slight effect upon the problem as a whole.

In some countries however these pre-existing relief systems of the unions have been utilized by government bodies as agencies through which to extend the principle of unemployment insurance at public expense. In 1901 the city of Ghent instituted what has come to be known as the *Ghent system* of unemployment relief, setting aside public funds for the payment of supplementary benefits proportioned to those of the trade unions. This system is administered through the machinery of the unions and is to a certain extent safeguarded against waste of public money by the requirement that no relief will be given except in consideration of a proportional payment out of union funds. The Ghent plan has been adopted by many of the principal industrial cities and minor governmental districts of Europe. It has two great weaknesses. In the first place, since the employers are not called upon to contribute to the relief of the unemployed, the system adds nothing to their inducement to keep their workmen employed. It does not of itself operate to reduce the unemployment evil. In the second place it can be applied only to the organized workers. It leaves out of account the unorganized, who are frequently the most necessitous.

Unemployment insurance in Great Britain. Great Britain was the first country to adopt a nation wide system of compulsory unemployment insurance and the only country to give such a system thorough trial. A law passed in 1912, applying to a few selected trades and some two million workers, set up an unemployment fund, to which the workers were compelled to contribute a certain amount per week while employed, the employers a certain amount for each worker on their pay rolls, and the govern-

ment the balance required to cover the pensions or "doles" guaranteed those out of work. The scope of the system was greatly extended during the years following its origin. By 1920 almost every wage-earner in Great Britain between the ages of 16 and 70 was covered by unemployment insurance. In 1924 some 12,000,000 workers, organized and unorganized, were contributing to the fund and entitled to its benefits; of these 8,330,000 were men, 3,430,000 women. At this date the weekly out-of-work benefit for men was $3.75, plus $1.20 for a dependent wife and 24 cents for each dependent child under fourteen. The workers contributed 18 cents per week each, the employer 20 cents for each workman on the payroll, and the government the balance required to maintain these benefits plus the cost of administering the scheme. Unemployed women and children were entitled to smaller benefits, and they and their employers made smaller contributions to the fund. The scheme attempts to reduce unemployment and to stabilize industry, by allowing a refund of one third their contribution to those employers who keep their men at work not less than forty-five weeks in the year.

Various safeguards are provided against abuse of the system. A waiting period of six days must elapse before the workman may draw an out-of-work benefit. While receiving the dole he must report daily to the local labor exchange and must accept any work which is offered provided it pays the normal rate of wages and is not in a concern where a strike is in progress. Striking workmen are not entitled to unemployment benefit, nor are those who leave their jobs without sufficient reason or are discharged for bad conduct. Fifteen weeks was at first the maximum period for which a workman could receive the dole. This was increased to twenty-six weeks in 1921. During the prolonged depression which set in about this time, the limit has frequently been removed entirely and the unemployed permitted to draw their benefits for an indefinite period.

The weakness of public unemployment insurance. Interesting evidence is furnished by the recent experience of Great Britain. Though unemployment is an assurable risk when judged by the

tests mentioned on a preceding page, it involves a large "moral hazard." This term, as used in the language of insurance, refers to the danger that the risk against which insurance is carried will be intentionally induced by the beneficiary of the insurance. Unfortunately there are some workers whose standards of living are so low and whose moral integrity so unreliable that they will gladly accept an unemployment benefit barely sufficient to maintain life if they can be exempt from all labor. Such men have no incentive to efficiency while at work; on the contrary they will deliberately make themselves unemployable by habits which, though they do not debar them from the benefit, will make every employer loath to retain them on his payroll. These parasites on the unemployment insurance system will drift about from industry to industry and from place to place, accepting work when obliged to do so in order to maintain their eligibility for the benefit, but hoping to be released from each new job as speedily as possible.

In times of depression, when there is much suffering among self-respecting workers, the safeguards against malingering are relaxed out of compassion for the great body of the unemployed. The shiftless and pauperized laborers find it easier to evade the stringent requirements of the law at such times, while the pervading hopelessness and distress tend to spread the poison of pauperism through the ranks of the laboring population. It is just at such times also that the drain of the insurance system upon the public revenues is at its height because involuntary unemployment is both widespread and protracted. To save the entire system from collapse it is imperative that every possible measure be applied to reduce the amount of unemployment and to increase the efficiency of the workers. But the system operates, when unemployment is at its height, to counteract efforts to solve the problem and make the working classes once more self-supporting.

This danger from the unemployment system did not appear in Great Britain until after the World War. At the end of its first five years of operation the insurance system was pronounced a success by most well-informed people in the country and had

aroused such interest in other countries that many governments were making plans for legislation along similar lines. It had been found that in normal times only two per cent of the unemployed were still out of work at the end of the fifteenth week. The burden on the government was not excessive; in 1920 there was a surplus of $100,000,000 in the fund. Nor had great difficulty been encountered in guarding against occasional instances of malingering. Since 1920 however the situation has changed entirely. Great Britain has passed through a protracted period of depression, with the number of continuously unemployed ranging between one and two millions. The weekly dole to this great army of idle people has caused a serious strain on public revenues at a time when the resources of the treasury were strained to the limit by costs resulting from the war. Since the war Great Britain has paid almost one and a half billion dollars in out-of-work benefits. In addition to its regular contribution, the government has been obliged to lend some $150,000,000 to the fund to keep it solvent. No doubt the weekly dole has been the salvation of millions of people during this time of great hardship, but there is reason to fear that the system has tended to aggravate the disease it was designed to remedy. The share of the cost placed on employers has burdened industries struggling to maintain themselves in a bad market; this acts as a deterrent to the creation of employment. There is evidence that prolonged distress has begun to break down the self-respect of the workers and to increase the number on the alert to abuse their privileges under the system. At the present time all political parties in Great Britain and most private observers express grave misgivings about the practicability of the policy of unemployment insurance.

Unemployment insurance through private action. A few years ago it appeared certain that the United States was on the point of following Great Britain's lead in the matter of compulsory unemployment insurance. Bills incorporating the major features of this system were introduced into the legislatures of Pennsylvania, Wisconsin, and Massachusetts, and commissions were studying the problem in California and New York. The recent

turn of events in Great Britain has put an end to this development in our country, at least for the immediate future. Many American employers however have voluntarily adopted plans of their own for the guarantee of employment or the indemnification of their workmen for losses incurred through unemployment. Few of these plans can properly be called insurance, since their continuance depends upon the willingness and ability of the employer to finance them. But it is possible that a sufficiently large number of successful experiments of this nature may lead to a new type of insurance adaptable to this great hazard of the wage-earner's life. One outstanding example of this type of plan is that of the Dennison Manufacturing Company. This company guarantees all its workmen against loss of earnings after the first half day of idleness; unemployed workers with dependents receive 80 per cent of their average wage, and those without dependents 60 per cent. The Dutchess Bleachery Company has a similar plan, though with different terms of compensation.

Another line of development in the United States takes the form of an unemployment agreement between organized labor and employers' associations. The entire clothing trade of Chicago is covered by such an agreement. By its terms the workers contribute one and a half per cent of their wages, the employers depositing an equal sum. This fund, made up of three per cent of the total pay roll, is drawn upon to pay idle workmen forty per cent of their wages after the first two weeks of idleness. The benefit in each individual case cannot exceed $20 per week nor extend for longer than five weeks continuously. A similar plan embraces the 35,000 clothing workers and the 1,500 employers in the New York market. The arrangement is much the same as in Chicago, except that the employers in New York contribute double the amounts paid by the workmen. In the Cleveland ladies' garment industry a variation of the same plan places the entire burden on the employers, who contribute seven and a half per cent of their pay rolls to an unemployment fund. This plan is so arranged that the employers receive rebates of their contributions proportioned to their success in providing continuous employ-

ment. In all the cases mentioned the chief motive which caused the employers to enter into the arrangement was their desire to stimulate the productivity of their workers. Fear of unemployment was known to be the most potent cause of restriction of output. During their brief periods of operation, these voluntary unemployment insurance plans have proved effective in accomplishing this major purpose. This in fact is their chief significance to the student of economic problems. If it can be proved that to guarantee the workers against loss through unemployment is good business, there can be no doubt that the self interest of business managers will sooner or later inspire them to find a solution for this great social problem.

CHAPTER LII

PROBLEMS OF THE WORKING DAY, CHILD AND WOMAN LABOR, AND INDUSTRIAL ACCIDENT

Modern aspects of the problem of the working day. During the past century there has been a consistent tendency toward a shorter workday for the average man. This tendency did not assert itself during the first decades of the Industrial Revolution. In the early factory system of England uncontrolled and unregulated competition for employment within the ranks of the newly created wage-earning population resulted in hours of work rising as high as fourteen per day. This was true also of the early factories in the United States. Such hours meant complete exhaustion of the productive energy of the men, women, and children who worked the machines. But this stage was short lived. A reaction soon set in which has left a record of almost continuous shortening of hours of labor down to the present day. This is shown in the following table.[1]

HOURS OF LABOR IN MECHANICAL AND MANUFACTURING ESTABLISHMENTS

Year	*Average Workday*
1840	11.4
1850	11.5
1860	11.0
1870	10.5
1880	10.3
1890	10.0
1900	9.8
1910	9.25
1920	9.0

At the present time hours longer than ten per day are recognized as exceptional, an anachronism lingering on in certain isolated

[1] Figures for 1840–1900 adapted from table given in Adams and Sumner, *Labor Problems*, p. 518; for 1910 and 1920 from figures published by United States Bureau of Labor Statistics.

industries. Enterprises in which trade unionism is strong have accepted the eight hour day or the forty-eight hour week as the standard, while in some few lines of occupation this has been reduced to the forty-four hour week, allowing a half holiday on Saturday. There is indeed evidence indicating that the movement toward the shorter day has not yet spent its force. That the modern wage-earner rebels against hours of labor which were accepted by former generations as a matter of course, that it has been found possible to grant his demand for a shorter day, that in so doing his real wages have not been reduced but actually increased — these are the aspects of the problem which require explanation.

Background of the movement toward shorter hours. The forces which have called into existence the pressure for the shorter day are not wholly economic in character. They spring in part from the general improvement in the intellectual and spiritual condition of the mass of mankind which has accompanied the material progress of the past century. Diffusion of education among the common people and the multiplication of their social and political activities create desires which can be satisfied only on condition that the average man enjoys a considerable period of freedom from the routine tasks of industrial employment. The services of libraries, museums, and amusement parks, provided free of charge by modern society, and the many forms of cheap entertainment which are familiar branches of business enterprise stimulate this desire for leisure.

But the demand for the shorter day is also in part a product of the peculiar economic conditions of the modern factory system. The truth is that the productive life of the wage-earner today is often a monotonous and joyless experience. In such a situation there is no stimulus to effort except that afforded by the wage rate and no immaterial benefits in the form of satisfaction from work well done. The control of productive enterprise is lodged in the hands of those who own the capital instruments or their representatives; the workmen have no share in the planning and direction and can feel none of the joy of achievement which comes to the

successful enterpriser. Finally fatigue increases progressively as the strain of effort is concentrated on a narrower range of muscular and mental faculties. Continuous labor of a certain duration, when it calls into play a large variety of the worker's powers, is not so exhausting nor so repellent as labor of the same duration performed by a limited set of muscles. Those conditions grow more pronounced with the development of the factory system and the progress of machine technique. It is now believed by those who forecast the future of industrial society that it will prove impossible to make the productive life of the average wage-earner anything but a dull and routine performance and that he must rely for self-expression upon ample hours of leisure spent outside the walls of the factory.

Relation of hours to wages. The desirability of the shorter day as tested by considerations of general moral and social welfare is granted by most people. But every attempt to reduce the hours of labor raises the question of what wages can be paid for the shorter service. When the workmen press for a reduction of hours, they invariably demand that the hourly rate of wages be raised so that their earnings will remain unchanged. In fact organized labor has a slogan which declares that a shortening of the workday of itself raises wages:

> "Whether you work by the piece or the day,
> Decreasing the hours increases the pay."

Of course this is merely one aspect of the "lump of labor theory."

On the other side are those who argue that every shortening of the workday must cause a decline of real wages, since the income of the worker is based upon his productivity. But this statement does not accord with the facts of experience. If we review the last seven decades we find that a decline of about two and a half hours in the average workday has been accompanied by a rise of real wages for all groups of workers. Of course there have been during this period other forces operating to the advantage of the laborers. The question still remains as to the limits within which the working day may be further shortened without causing a loss of real wages.

Reduction of hours and increased productivity. Within limits a reduction of hours has been found to increase the productivity of labor. It has been proved by the experience of many different enterprises in Europe and America and established as a rule of wide applicability by researches carried on by the British Government and by various bureaus of investigation in this country that a shortening of the more extreme workday — a fall of hours from twelve to eleven or ten — results in an actual increase in the total output of labor.[1] To be sure, this effect of shortening the long workday is not uniform for all grades of workers. It is most noticeable in the case of those occupations which demand continuous mental alertness combined with manual dexterity and least so of the automatic types, such as the tending of machines whose pace is governed by the flow of steam or electrical power. But even the most unskilled labor shows some response to a reduction of an extremely long day by an increase of output.

There are many reasons for the increase in productive efficiency which results from a reduction of the more extreme workday. When the length of the workday is excessive, the period of rest is not sufficient to counteract the effects of fatigue. In extreme cases accumulated fatigue predisposes the workman to sickness which results in lost time, and where this consequence is avoided the workman tends to be slack in attendance through voluntary idleness. It has been found that the rate of accident varies directly with the length of the workday, reaching a maximum during the late hours, and also that it rises with the days of the week, thus disclosing another waste due to fatigue. Waste of materials, in the form of defective output, waste of power, and necessity of more labor of superintendence, through increased slackness of the workers, also result from protracted hours of labor. A saving of these wastes can be effected by shortening the hours of work to avoid the worst effects of fatigue. It is probable that the ten hour day is more productive at the present time for all grades of

[1] The evidence in support of this conclusion is summarized in the *International Labor Review* for April and November, 1924. See also Vernon, *Industrial Fatigue and Efficiency*, London, 1921.

labor than is the twelve hour day and that for the more skilled groups an eight hour day is at least as productive as ten hours.

Limits to reduction. But there are certainly limits beyond which reduction of hours cannot be carried without diminishing the total output of the workers and consequently causing a fall of real wages. It is difficult to determine these limits with precision, since they differ for each of the various grades of labor comprised in the working force of the typical industrial establishment. But to disregard them when reducing the hours of work is to diminish the material prosperity of the workers and of society as a whole. On this point we have evidence in the trend of real wages as compared with the length of the workday during the first two decades of the present century. From 1900 to 1920 the average workday was shortened by approximately one hour; weekly earnings measured in money rose 80 per cent, but the advancing cost of living more than offset this rise of money wages, resulting in a fall of real wages of a little more than 20 per cent. This period, it is true, was so upset by a fluctuating price level and by a violent swing of the business cycle that we cannot accept the testimony of these statistics with complete confidence. Nevertheless it seems apparent that the shorter workday was gained by many of the workers at a sacrifice in terms of material welfare. And this conclusion is strengthened by investigations which have been made of the effects of the shorter day on the productivity of certain grades of labor. It has been demonstrated that in many cases the gain in hourly productivity resulting from a substitution of the eight hour for the ten hour day was not sufficient to maintain the former level of output.

Effects of invention. Thus far our analysis of the economic effects of the shorter workday has been carried out on the assumption of a fixed stage of industrial technique, leading to the conclusion that there is a point beyond which the workday can be shortened only at the expense of a reduced product and a lower real wage. But in a society subject to rapid changes in industrial processes, as is ours, each new invention tends to increase the product of the average worker in a given length of time. And

when technical progress is comprehensive, applying in general throughout the whole field of industry, its ultimate effect must be to expand the material product of all groups of workers in a workday of given length.

When real wages have thus risen high enough to establish a relatively comfortable standard of living for the higher groups of workers, each successive advance in industrial technique presents these workers with a choice between two alternative ways in which to take their share in the general improvement. They may elect either to receive their former rate of real income and to take their gains in shorter hours of work or working the same number of hours as formerly, to raise their standard of material comfort. This choice has been presented to the wage-earning population, without their knowledge, in the course of the rapid industrial evolution of the last three quarters of a century. In many cases their actual choice has taken a middle ground between the two extremes. As their bargaining power has increased, the workers have pressed for a shorter workday and for a simultaneous increase of wages. That they have won both concessions is to be attributed in part to the positive increase of productivity resulting from reduction of long hours which we have discussed, in part also to the fact that these demands were made at a time when mechanical progress was working in the direction of increasing the hourly productivity of labor. For some groups of workers the shortening of the workday has maintained their productivity at its maximum; for others it has resulted in an output below the maximum of which they are capable on the present stage of the industrial arts, but higher than was possible on earlier stages. At the present time it would appear that further reductions in the workday, if they are not to be attended by impoverishment for the average man, must generally wait upon further advance in the arts of industry.

Modern aspects of child and woman labor. The problem of woman and child labor, like that of the length of the workday, can be called a modern problem only if we qualify our terms. The lot of the women and children of the poorer classes in society

has always been one of toil, and it is probably true that both in respect of the conditions under which they work and of the rewards for their labor the modern world deals less hardly with those more defenseless of its members than did society in its earlier stages. But it may be said nevertheless that the labor of women and children is a modern problem, not because it is more extensive or more degrading than formerly, but because its effects are different.

To grasp the significance of such a problem as this, at any stage of societal development, it must be studied in its relation to the whole complex of customs, institutions, and ideals which comprise the social structure of the time. It is because this relationship has been changed by the revolution of the social order which produced the modern industrial world that we are correct in speaking of the labor of women and children as a modern problem. The change has been due in part to an alteration in the ideals and principles which govern our political and social life. The tests of enduring social welfare accepted in a democracy may be quite different from those which obtain in an aristocratic order. Conditions which deny to children opportunities for wholesome physical development and which interfere with the normal functions of motherhood or contribute to the spread of vice among unmarried women are recognized as constituting a problem of obvious social import. It is from the standpoint of its relation to the general social welfare and to the ideals of the modern world that we must appraise the significance of child and woman labor.

But there is another way in which this problem differs from that of an earlier social order; namely, in its effects upon the life conditions of male workers. Before the Industrial Revolution the labor of women and children was supplementary to, not in competition with, that of men. With the perfection of machine technique however processes were so far simplified that jobs formerly monopolized by men could be filled by children or by women who possessed no considerable skill or physical strength. At once craft lines began to dissolve; cheap labor began to compete for employment with the skilled worker; women and children were in competition with men. This threatened a lowering of the stand-

ard of life and an undermining of the security of the male workers. With the progress of the Industrial Revolution this condition spread from one industry or process to another until today it constitutes one of the major forces shaping the policy of the male workers.

Extent of child labor. Government statistics of child labor do not exist for the period prior to 1870. The census of that year reported 739,164 children between the ages of 10 and 15 at work. In 1880 the number had risen to 1,118,356; in 1900 it was 1,750,178; in 1910, 1,990,225. The decade from 1910 to 1920 is the only one in forty years when the number of child workers fell; in 1920, there were 1,060,858 children of this age group at work. It is probable that this decrease was due to the fact that the federal child labor tax law, which virtually prohibited the employment of children below the age of 16 in mining and manufacturing establishments, was in effect when this last census was taken. The law was later declared unconstitutional, and it is estimated that the number of children at work has risen again to approximately the two million mark. Of this total, about 65 per cent are employed in agriculture, 20 per cent in manufacturing, and the remainder in trade, clerical occupations, and domestic service.

Effects of child labor. The evils of child labor do not require elaboration. The injurious moral effects which flow from an imperfect education, a denial of the period of play and recreation indispensable to mental and physical health, and contact with an environment which is often contaminating are apparent to all. There are as well certain economic effects which react unfavorably on the material welfare of society at large. The wage-earning child is restricted in his choice of occupation to lines in which there is no hope of considerable advancement in earning power. Even when he is not physically crippled by over-strain or by abbreviation of the normal period of rest, he is denied opportunity to acquire the skill demanded of those in the higher wage groups. At maturity he is usually caught in a "blind alley" occupation or recruited to the ranks of the already overcrowded group of unskilled manual labor. This involves a loss not only to the

victim of child labor but to all the rest of society's members. For, as we have seen in connection with our study of differences in wages, the total income of society would be increased if it were possible to transfer men from the relatively overcrowded groups whose marginal productivity is low to the less crowded occupations where marginal productivity is higher. Any device which hinders this transfer not only tends to perpetuate the inequalities of income as between the different classes in society but also prevents a possible net increase in the total income fund.

Child labor is one of those social problems which seem unable to work out their own correction. In some forms of occupation children can be used in place of men at a saving in cost to the employer. Since they do not rely wholly upon their earnings for support, their wages are lower than the minimum at which it would be possible to maintain an adequate supply of self-supporting adult workers. In this there is an inducement, heightened by the pressure of competition, for the enterpriser to employ children wherever possible. This tendency is not counteracted by opposition from the parents. In very many cases, either through the pressure of their own poverty or through blindness or indifference to the welfare of the child, parents willingly consent to industrial employment for their children. The children themselves are obviously unable to resist. Nor do they show a disposition to do so. Quite naturally they rebel against the discipline of the school and, tempted by the prospect of a position of independence, are glad to engage in wage labor at the sacrifice of their future. When all parties to the contract coöperate thus willingly to increase the spread of child labor, the only remedy is that of restrictive legislation.

Legal regulation of child labor. Early attempts to regulate child labor by law had to fight their way against the opposition of prevailing individualism. The first Act of Great Britain, passed in 1802, attempted no more than the prohibition of the labor in factories of pauper children below the age of nine years or for more than twelve hours a day or at night. The mere mention of these limits is enough to show the gravity of the problem of child labor

during the early stages of the Industrial Revolution. Opposition even to mild limitation is shown in the series of amending and supplementary acts which were necessary before anything like thorough regulation of child labor could be accomplished. Not until 1833 were all textile mills brought under legal regulation. The law of that year prohibited the labor of children under the age of nine and provided that children between the ages of nine and thirteen should work only eight hours a day and those between thirteen and eighteen only twelve hours, with no employment at night. The factory acts were extended to all large industries in 1864 and to smaller concerns in 1867. In the meantime an Act of 1842 had prohibited the labor of children under ten years of age in mines, which limit was later raised to thirteen years.

Without following in detail the development of this legislation, we may note the progress of thought on the subject by comparing with these early standards the terms of Great Britain's latest statute. By the law of 1920 fourteen years is made the minimum age for industrial employment for all children; the length of the workday for young persons between sixteen and eighteen years is limited to eight hours; no work is allowed at night for any wage-earner below the age of eighteen. Other European nations of advanced industrial civilization have approximately the same standards. These laws are sanctioned by heavy penalties and administered by an efficient system of government inspection which makes evasion extremely difficult.

In America. The regulation of child labor in the United States lies in the hands of the states, and practically every state has a law on the subject. Most of these laws agree in fixing the minimum age at fourteen years. Thirty-eight states prohibit night work for children under sixteen; twenty-six debar children from mines and quarries. Some thirty of the states have fixed a limit of eight hours to the working day of children under sixteen, while others permit a working day of nine, ten, or eleven hours. The laws do not apply to work in agriculture or domestic service, and only a few of the states have attempted to regulate the street trades of children. The practical effect of these laws varies, not

only because of their different standards but also because of the wide difference in the rigor with which they are administered in different states. In general the Southern states, where the factory system is of late development and where a necessitous population with low living standards creates a large supply of child workers, have the lower standards of regulation.

Lack of uniformity has been found to affect injuriously those units of a competitive industry which are located in the states having the more rigid standards of regulation, and this fact has tended to arouse opposition to child labor laws in these advanced states. Consequently various attempts have been made to bestow on the Federal Government power to impose uniform standards upon the country as a whole. A federal law passed in 1916 attempted to prohibit the labor of children under fourteen, to limit the workday of those from fourteen to sixteen, and to prohibit the labor of children at night, by excluding the products of industries violating these standards from interstate commerce. When this law was declared unconstitutional in 1918, Congress passed another statute embodying the same standards and attempting to regulate child labor by applying prohibitive taxes to the concerns violating the standards. This law also was declared unconstitutional, in 1922. The last attempt to give the Federal Government authority over child labor took the form of an amendment to the constitution adopted by Congress in 1924. The effect of this amendment would have been to bestow on Congress power to prohibit or regulate the labor of children under eighteen years of age. However ratification of the amendment by the necessary majority of the states has not been obtained.

Extent and character of industrial labor of women in the United States. Though the number of women workers in industry in this country represents a small minority of the entire labor supply, this group of workers has been steadily increasing both absolutely and relatively during recent years. The census of 1880 recorded 2,647,000 female workers over ten years of age gainfully employed, this being about fifteen per cent of all the women in the country. By 1900 the figure had risen to 5,320,000, or nineteen per cent.

In 1910 there were over 8,000,000 women employed, twenty-three per cent of the female population. The next decade showed a slower rate of increase; in 1920 the number of women workers had risen to 8,500,000, but the proportion to the female population had fallen to slightly above twenty-one per cent. It is probable however that this decline in the relative importance of the women workers is to be explained as the result of a change in the method of compiling the census figures and that it does not indicate a permanent tendency in the direction of reducing the importance of this factor in our labor supply. During this entire period, from 1880 to 1920, the number of women wage-earners has increased more rapidly than the increase in the total working population. The ratio of male workers to female has fallen slightly since 1880. In 1920 women comprised more than one fifth of the entire wage-earning population.

On the whole women enter lines of occupation where they are not brought into direct competition with men. There is a marked tendency for certain forms of labor to be monopolized by men and for others to fall largely into the hands of women. This is shown by the following table, which is based upon an analysis of the census of 1920.[1]

PER CENT DISTRIBUTION BY SEX OF LABOR IN THE GENERAL DIVISIONS OF OCCUPATION

Occupation	*Male*	*Female*
All occupations	79.5	20.5
Agriculture, Forestry, Animal Husbandry	90.1	9.9
Mining	99.7	.3
Manufacturing	84.9	15.1
Transportation	93.0	7.0
Trade	84.3	15.7
Public Service	97.2	2.8
Professional Service	52.6	47.4
Domestic and Personal Service	35.8	64.2
Clerical Occupations	54.4	45.6

As will be seen, the competition between men and women is of chief importance in those occupations which we call personal service — professional, domestic, and clerical occupations. Outside

[1] Taken from Leo Wolman, *Growth of Trade Unions*, 1880–1923, p. 100.

this group the labor of women has grown to significant dimensions chiefly in manufacturing and in trade, where about fifteen per cent of all those employed are women. A more refined classification of these two occupational groups would show that woman labor concentrates in certain forms of manufacturing, such as textiles and clothing manufacturing and in certain subdivisions of the group which is called trade.

Causes of the increase in woman labor. To some extent the greater employment of women is due to the progress of machine technique, to which reference has been made at other points in our study. Quickness and deftness of touch take the place of muscular strength and endurance as qualifications in the labor supply of those industries which are equipped with automatic power-driven machinery. This makes it possible to substitute the labor of women or children for that of men.

But the movement is not wholly accounted for by reference to this factor. Mechanical equipment is of chief importance only in manufacturing, where, as we have seen, the labor of women has not risen to such formidable proportions as in other lines of occupation. Moreover the possibility of adapting the job to the characteristics of the woman worker does not of itself suffice to explain why the women of the modern world are more willing than in former years to seek employment as wage-earners. For an explanation of this fact we must look to tendencies of broad social significance in the customs and habits of our population. The postponement of marriage which is so noticeable in the modern world as compared with that of preceding generations, the removal of customary restraints which formerly circumscribed the status of women in the home and neighborhood, and the increase in amount and change in character of the education offered to women are all forces working toward an increase in the disposition of women to seek positions of economic independence.

Results of women's increased part in industry. We shall not undertake to make a final appraisal of the net result of this feature of our industrial system in terms of social welfare. It should be noted however that the tendency toward increased employment

of women, founded as it is on such fundamental forces in our social order as have been suggested, is probably irresistible. There are unquestionable gains resulting from the influx of women into industrial occupations. For the women themselves, this change in the conditions of their lives has meant escape from the inanity and futility of woman's existence in the Victorian age and an opportunity for self-expression and self-development which must contribute to a richer and more significant life. Society has gained materially, through the substitution of productive for meaningless tasks for its unmarried women, and spiritually, through the release of energies formerly thwarted by the bonds of custom and tradition. But these benefits have not been unmixed with other consequences. Certain grave moral problems arise in connection with woman labor, which society cannot lightly disregard. On economic grounds also account must be taken of certain peculiar results, arising chiefly from the bargaining weakness of the woman wage-earner.

The problem of equal pay. The general attitude of the male worker toward the labor of women is based squarely on his desire to protect himself from what he considers to be unfair competition. We have seen that this competition takes place chiefly within the range of certain specific industries. Within these industries or certain subdivisions thereof, the proportion of women workers may be large enough to determine the conditions on which male workers can find employment. The device on which the men rely for protection against the competition of women is summed up in the slogan of organized labor: "Equal pay for equal work."

Two very different meanings may be attached to this demand that the wage rate for women be the same as that for men. In the first place it may mean that the payment for a given product should not be less for women, solely because they are women, than it is for men. Since women are admittedly less productive in most occupations than men, the equal pay slogan in this sense would result in lower day wages for women. The second meaning resolves the slogan into a demand that the daily or weekly wage of women be the same as that of men, regardless of the relative

productivity of the two groups of workers. If applied in this sense the result would be higher wages in relation to productivity for women than for men. As advanced by most of the labor unions, the demand for equal pay has the second of these meanings. Its intention is not to raise the wages of women, but to exclude them from employment. If employers are required to pay the same wage to two groups of different productivity, it is obvious that they will not be inclined to employ the less efficient workers at all. On economic grounds there is little to be said in defense of the doctrine of equal wages when applied in this sense.

Underpayment of women. Even if we limit the meaning of the equal pay doctrine to that sense in which it asserts the justice of paying equal wages in proportion to productivity, it remains true that this is not the practice of the business world. On the contrary the woman worker has been consistently underpaid as compared with the male worker. This matter was investigated in the industries of Great Britain during the World War, in response to complaints from women who felt they were being victimized by the wage policy of employers. The investigation disclosed a widespread practice among employers of paying a lower wage to women than to men in exchange for the same service. Even under the piece rate system, in which wages are based upon the volume of the worker's output, some rates were found to be as much as fifty per cent lower for women than for men on the same job. Certain factories were shown to have lowered their rates when women were employed in the places of men, for no other apparent reason than that the new employees were women. Similar studies in America have established the same fact. Even allowing for all concealed factors which might explain the discrepancy in wage rates on the ground of the inferior productivity of women, the conclusion cannot be escaped that it is customary to pay less than equal wages for women.

This systematic underpayment of women cannot be accounted for except on the ground of their weaker bargaining power. As compared with men, women are notably weak in asserting their claims against the employer. To their naturally timid and non-

aggressive disposition are added handicaps which arise from their social position and their economic environment. We have seen that they are concentrated for the most part in certain lines of occupation. Here they meet severe competition from a fluctuating group of women laborers who seek part time or temporary employment as a supplement to their labor at home. Again they are even more immobile than the men; it is difficult for them to play off one employer against another when this involves a change of residence. In many cases also they are not self-supporting and make no pretense of demanding a "living wage"; the minimum wage which they will accept is lower than that of men, because they are partially subsidized by the male wage-earners of their families. Those young women who look upon their employment as a temporary expedient to fill in the time before their marriage are not always intent upon exacting the highest possible price for their labor. Finally they are unorganized and to a large extent unorganizable; their bargains are made as individuals, and they labor under all the handicaps which the isolated worker must bear in the modern labor market.

Legal regulation of woman labor. Aside from the minimum wage laws of our states, which were discussed in another place, laws bearing on the labor of women have been devoted primarily to setting a maximum limit on their hours. The constitutionality of these laws, which for many years was in question, was decided in 1908 by the action of the Supreme Court in upholding the ten hour law of Oregon. This action gave impetus to the spread of legislation limiting the hours of women, so that today (1925) only four of the states have no laws on the subject. The limit varies between eight hours and ten in most of the states; though there are four which allow a workday in excess of ten hours. In general women are excluded from mines and other unusually hazardous occupations. Little is done to regulate night work or home work. The laws specifically exempt farm work and domestic work from regulation.

The waste of industrial accident. One of the principal costs involved in our modern system of industry is due to the temporary

or permanent destruction of the productive power of the workers through injuries resulting from the normal hazards of their occupations. Only within recent years has serious attempt been made to estimate the magnitude of this cost. The official statistics of work accidents in the United States are even now so unreliable and incomplete that any statement of the total loss can be no more than an estimate. In some states no records at all are kept. Certain industries have accurate records of the fatalities resulting from the hazards of the occupation, though the less serious accidents which result in temporary disability are allowed to go unrecorded. In the mining of bituminous coal for example we know that about two thousand men are killed each year, but we have no accurate figure for the non-fatal accidents except such as can be inferred from the known fact that in the state of Pennsylvania alone over 28,000 such accidents occur in the course of a year, resulting in a loss of over a million days of labor. Estimating from such statistics as we possess, a conservative statement of the fatalities in all branches of industry would place the number at about 35,000 each year. The non-fatal accidents are estimated at upwards of 2,000,000, of which 500,000 result in disability lasting more than a week.[1] Here is obviously a waste of great magnitude.

Attempts to reduce the waste of industry. In recent years strenuous efforts have been made to reduce this waste. Factory acts and compensation laws have made it obligatory or profitable for individual owners of business enterprise to safeguard the lives and health of their employees. Groups of employers in the major lines of industry have coöperated through their trade associations to perfect devices for the prevention of accident and to urge their adoption upon the more indifferent individuals among their number. In the United States this effort has given rise to a widespread "safety first" movement, which enlists the services of experts and adopts many ingenious devices to make its program effective. Similar movements have been under way for decades in other countries, reaching their point of highest efficiency per-

[1] See Downey, *History of Work Accident Indemnity*, p. 1.

haps in Germany, which leads the world in accident prevention. There can be no doubt that these concerted efforts have prevented a vast number of accidents which otherwise would have occurred.

The unavoidable loss of accident. But the most impressive result of such efforts has been to demonstrate that this cost is on the whole unavoidable — an inevitable feature of our industrial system. After every possible precaution has been taken to prevent them, work accidents show a tendency to increase, both in total number and relatively to the laboring population. This is made clear in the following words of one of the greatest authorities on the subject:[1]

"The ugly fact is that work accidents, in the main, are due to causes inherent in mechanical industry on the one hand and in the hereditary traits of human character on the other hand. In the first place, a high degree of hazard inheres in present day methods of production. Modern technology makes use of the most subtle and resistless forces of nature — forces whose powers of destruction when they escape control are fully commensurate with their beneficent potency when kept in command. Moreover, these forces operate not the simple hand tools of other days, but a maze of complicated machinery which the individual workman can neither comprehend nor control, but to the movements of which his own motions must closely conform in rate, range, and direction. Nor is the worker's danger confined to the task in which he is himself engaged, nor to the appliances within his vision. A multitude of separate operations are combined within one comprehensive mechanical process, the successful consummation of which requires the coöperation of thousands of operatives and of countless pieces of apparatus in such close interdependence that a hidden defect of even a minor part, or a momentary lapse of memory or attention by a single individual may imperil the lives of hundreds. A tower man misinterprets an order or a brittle rail gives way, and a train loaded with human freight dashes to destruction. A miner tamps his 'shot' with slack, and dust explosion wipes

[1] *Op. cit.*, pp. 3–4.

out a score of lives. A steel beam yields to a pressure that it was calculated to bear, and a rising skyscraper collapses in consequence, burying a small army of workmen in the ruins.

" In the second place, human nature, inherited from generations that knew not the machine, is imperfectly fitted for the strain put upon it by mechanical industry. Safely to perform their work, the operatives of a modern mill or railway should think consistently in terms of those mechanical laws to which alone present day industrial processes are amenable. They should respond automatically to the most varied mechanical exigencies and should be as insensible to fatigue and as unvarying in behavior as the machines they operate.

" Manifestly these are qualities which normal human beings do not possess in anything like the requisite degree. The common man is neither an automaton nor an animated slide rule. His movements fall into a natural rhythm, indeed, but the beat is both less rapid and more irregular than the rhythm of most machines — with the consequence that he fails to remove his hand before the die descends or allows himself to be struck by the recoiling lever. It requires an appreciable time for the red light or the warning gong to penetrate his consciousness, and his response is apt to be tardy or in the wrong direction. Fatigue, also, overcomes him, slowing his movements, lengthening his reaction time, and diminishing his muscular accuracy — thereby trebly enhancing his liability to accident.

" The machine technology, in fact, covers so small a fraction of the life history of mankind that its discipline has not yet produced a mechanically standardized race, even in those communities and classes that are industrially most advanced. And so there is a great number of work injuries due to the ' negligence of the injured workman ' — due, that is to say, to the shortcomings of human nature as measured by the standards of the mechanician. This maladjustment is aggravated by the never-ceasing extension of machine methods to new fields of industry, and the continued influx of children, women, and untrained peasants into mechanical employments. Accordingly, the proportion of accidents attrib-

utable to want of knowledge, skill, strength, or care on the part of operatives appears everywhere to be increasing."

It is obvious that all work accidents result in net social loss. Even disregarding the suffering of the victims and their families, society cannot but be injured by this continuous destruction and crippling of the units of an essential productive instrument. When a laborer is thrown into idleness society will be deprived irrevocably of that contribution to its income fund which he would have made. Since accidents cannot be wholly avoided, the only effective remedy lies along two lines of development: reduction of the number of work accidents by applying the resources of science to their prevention and, when accident does occur, rehabilitation of the workman as speedily and completely as possible so that his productive energy may be placed once more at the service of society.

Distribution of the loss. Though social loss from work accidents is inevitable, arrangements can be made which affect its distribution as between different individuals in society. Such arrangements do more than raise the question of justice. It has been found that different methods of apportioning the loss from work accidents have different effects upon the number of accidents and upon the magnitude of the social loss involved in them.

The burden must fall in the first instance upon the injured workman and his family. Where the rate of accident differs strikingly as between different occupations, this hazard will be one of the factors in the distribution of the labor supply. If the workers were entirely free to choose their lines of occupation and if they made this choice on the basis of cool calculation of costs and income, the more hazardous forms of employment would be relatively undermanned and the wage rate would be relatively higher than in occupations which were less hazardous. In this way compensation for the risk would be worked out by the forces of demand and supply, and the prices of the different kinds of goods would carry the cost on to the consumer. It is probable that a partial adjustment was made in this way before society undertook to distribute the risk by collective action. But we have seen that

the distribution of the labor supply is not controlled principally by rational choice but rather by conditions of birth and tradition, which result in a fairly rigid classification of the working population into non-competing groups. The supply of labor in the hazardous occupations is not materially limited, as compared with other lines of work, by the risk of accident, and the wage rate, though it may contain a partial compensation for this factor of cost, does not suffice to lift the burden entirely from the workman.

Damages under the common law. Before the passage of the compensation acts, the injured workman in the United States was entitled to sue the employer for the recovery of his loss. In the courts the justice of his claim was tested by the principles the common law defining the liability of employers. But these common law principles were evolved at an earlier stage of industrial development to govern the simpler relations of master and servant, and their maladjustment to the conditions of the modern world placed serious obstacles in the way of recovery by the workmen. Stated briefly, the court attempted to discover who was to blame for the accident; if it could be shown that the fault was wholly the employer's, the workman was awarded damages; but if it appeared that the workman himself or someone else shared the blame with the employer, no matter in how slight a degree, the case was dismissed without compensation. This was the essence of the common law *doctrine of negligence*, which in all Anglo-Saxon countries determined the employer's liability to indemnify his workmen in case of accident in the absence of legislation to the contrary.

The difficulty under which the victim of accident labored in his attempt to shift the burden of loss will be apparent when we consider the defenses to which the employer could resort in answer to the workman's suit. In the first place he could assert that the accident was due at least in part to *contributory negligence* on the part of the victim. The inability of the human organism to adopt its responses perfectly to the rhythm of the machine, coupled with the complicating effects of fatigue, which we have already noted,

make it probable that most accidents are due at least in part to some mistake of the victim.

Failing proof of contributory negligence, it might be shown that the accident had been due to some fault of a fellow employee. The increasing interdependency of the personnel of the modern factory and the responsibility borne by those workmen who control powerful machinery which acts as a constant menace to the lives and limbs of their associates makes the *negligence of the fellow servant* an important factor among the causes of accident. If any element of blame could be attributed to a fellow employee, the employer was released from liability.

If these defenses failed, the employer might yet show that the accident had been due to a normal hazard of the business, of which the workman was aware when accepting employment. Under the doctrine of *assumption of risk*, the victim of accident had no claim against the employer for injuries resulting from these normal and known hazards. Considering how little choice the unemployed laborer has in the matter of accepting the job which offers, it is clear that this defense would be in many cases effective. It is not surprising that the workman was rarely successful in recovering indemnity for his injury.

Compensation laws. Modern compensation laws require the employer to indemnify the injured workman for his loss in wages and medical care and to compensate the victim's family in case of fatal accident. These laws are based on a line of reasoning quite different from that which gave support to the common law doctrine of negligence. They give up all attempt to assess the blame for accident, recognizing that in our modern system of industry destruction of life and limb are inevitable normal costs not attributable to human fault. The financial burden involved in compensation for accident and death is accordingly placed solely on the employer. It is expected that this cost will become part of the enterpriser's expenses and enter into the price of the product, thus finding its way to the consumer, upon whom in the last analysis the burden should fall.

One of the greatest virtues of the policy of holding the employer

financially responsible for accident is that it operates to reduce the number of accidents. By making the hazards of occupation expensive to those who control industry, the compensation laws create an incentive in the group which alone has the power to prevent them. The "safety first" movement and similar campaigns for the reduction of accident sprang up after society had determined to hold the employer responsible for the cost of injury received in the course of employment.

History of compensation laws and insurance. Germany led the way in legislation requiring the compensation of injured workmen. A law of 1884 gave effect to the principle that all employers should be liable to their workmen for partial compensation for wages lost through injury (two thirds of the former wage is paid to those permanently disabled) and to the dependents of the workmen for the loss of earnings resulting from fatal accident. Industrial accident is an insurable risk. The German system worked out a basis of insurance within each major group of industries and compelled the employers in each of these groups to form mutual insurance companies, to which the different business concerns paid premiums graded according to their accident rates. The whole system is supervised by public officials. Its efficient management soon led to a declining accident rate in German enterprises, which has placed Germany in the forefront of modern nations as regards this waste of industry. In 1897 Great Britain passed a comprehensive act requiring the compensation of injured workmen. This act makes no attempt at compulsory insurance, merely placing on the employer the liability of paying one half the former wage in case of non-fatal accident and a lump sum roughly equivalent to three years earnings in case of death. As a practical matter British employers almost invariably insure their liability in some private insurance company, though not required to do so. France and Belgium have similar systems.

The United States was the last of the great industrial nations to adopt the principle of employers' liability for accident. The first state laws on this subject were declared unconstitutional, and it was not until 1909 that New Jersey succeeded in setting

up a system of compensation which was acceptable to the courts. From this date down to the present the spread of workmens' compensation laws has been rapid. By 1924 forty-five of the states and territories had passed laws on the subject, and the Federal Government has adopted a model system for the protection of its employees. Arkansas, Mississippi, Florida, North Carolina, and South Carolina were the only commonwealths in which the old common law doctrine of employers' liability was still maintained. The laws vary greatly in detail. The most advanced entitle the workman to a pension of two thirds his former wages in case of permanent disability and provide a similar pension for his dependents in case of death. In the majority of cases the employers are required to insure their liability, and the law frequently sets up a state insurance fund under public administration for this purpose. About a dozen of the laws, including that of the Federal Government, cover occupational diseases as well as accident.

CHAPTER LIII

THE ORGANIZED LABOR MOVEMENT

Foundation of the labor movement. In beginning our study of the organized labor movement, we have first to seek an explanation of the forces which have brought it into being. Here it is necessary to rid our minds at the outset of any disposition to view the labor union as an artificial growth, the work of a few labor leaders who have succeeded in persuading an apathetic mass of wage-earners to join the ranks. It is far too widespread a phenomenon to be accounted for by so simple an explanation. Labor unions seem to spring up inevitably in the path of the factory system as it spreads throughout the modern industrial world. From its first appearance in England, the birthplace of the factory system, we may trace the development of the labor movement in all the countries of Europe and America in rough correlation with the increasing industrialization of these countries. Sixty years ago Japan was a feudal country, in which labor unions were wholly unknown, although they flourished at that date in the industrial nations of Europe. During the intervening period Japan has been passing with great rapidity through an industrial revolution, until today the factory system is firmly established as a factor in her economic life. In the course of this evolution we may observe the origin of an organized labor movement, which discloses a tendency to follow the same lines of development as that of the older industrial countries. So it is in those parts of the world where the advent of the factory system is a still more recent phenomenon. In our own day India and Egypt are beginning to take on the industrial culture of Western Europe, and here too we note the beginnings of labor organization. According to the most reliable information available on this subject the labor movement through-

out the world as a whole numbered in 1920 48,000,000 members.[1] Of these, 34,000,000, or 82 per cent were to be found in European countries and seven and a half millions, or an additional 15 per cent in the countries of North America. The four great industrial countries, Germany, Great Britain, the United States, and France, contained 28,000,000 unionized wage-earners or about 67 per cent of the total membership of the world.

We must conclude that this movement is to be understood only as an attempt on the part of great numbers of men to adjust themselves to certain conditions of life created by the factory system. These conditions have been suggested by our preceding study and need only be summarized here. In the first place the factory system tends to separate the worker from ownership in the capital instruments and the land with which he works. This destroys the position of the laborer as an independent enterpriser and reduces the mass of men in society to something like a uniform economic status, welding them gradually into a distinct social class. Secondly it concentrates the workers in close contact with each other upon a limited geographical area. The discipline of the factory, by accustoming the wage-earners to a system of regimentation in which conscious coöperation is carried out under orders of management, provides conditions which conduce to voluntary group conduct. In the third place the factory system released certain unpleasant or injurious economic forces which affect the workers in common and against which they are as individuals powerless to protect themselves. These forces have been sketched in preceding chapters — unemployment, low wages, monotonous and hazardous forms of work, the labor of women and children — all problems whose solution is vital to the security and contentment of the mass of wage-earners. The existence of these common problems supplies the motive for collective action among workers, as the regimentation of the factory creates the condition which facilitates such action.

We must view the modern labor movement as the product of these conditions. It did not exist in earlier stages of industrial

[1] *Monthly Labor Review*, U. S. Bureau of Labor Statistics, Jan. 1922.

evolution; the medieval guilds, as we have seen, were quite different in form and purpose. Nor do labor unions flourish today among groups of workers whose life conditions are not affected by the factory system. In countries like England and the United States, where the movement is firmly established, there exist great numbers of workers on the farms, in the retail trade, or in other positions of semi-independence, among whom labor unions have been unable to gain a permanent foothold.

Origin of the labor movement in Great Britain. While space does not permit us to trace in detail the rise of the modern labor movement, some reference to its history is required in order that we may be in position to appraise its importance and to understand its peculiar features of structure and policy. Unions of workers, organized with the intent of controlling wages and conditions of employment, began to attract public attention in England before the opening of the nineteenth century. These were at first isolated trade or craft groups — workers engaged in the same skilled occupation — whose predominant motive was to secure the enforcement of the customary and legal guarantees as to wages and the right to work which had come down from the medieval period of England's industrial life. This attempt to perpetuate the conditions of an earlier economic system failed; economic forces generated by the factory system swept away the monopolistic rights of the skilled craftsmen and rendered obsolete the apprenticeship rules upon which these rights were based. The old laws fixing wages and regulating employment were repealed, and a public policy was established which accepted as its basis the principles of individualism and competition which were becoming dominant in the economic life of the country. In harmony with this newer viewpoint, all unions of workers for collective economic action were declared illegal by laws passed in 1799 and 1800. The following quarter century witnessed a rigorous repression of those illegal combinations of workers which persisted in forming despite the severity of the law. During this period the factory system was making rapid headway in England. Competition among the wage workers became increasingly bitter as the progress of ma-

chine technique undermined the security of the skilled trades and made possible the employment of women and children in processes formerly monopolized by men. Wages fell, and hours of labor and conditions of sanitation and safety in factories, mines, and tenements became intolerable. In self-protection the male workers in the different industries united to resist these destructive tendencies, directing their efforts toward an increase of their bargaining power in competition with their employers rather than toward a reëstablishment of medieval conditions. Fines, imprisonment, expulsion from the country — punishments visited by the law upon the leaders of these unions — did not avail to crush the movement. When the Combination Acts of 1799–1800 were repealed in 1824, the law merely took cognizance of the fact that organization among the workers was a natural feature of the new type of industrial society.

Conflict with the law. From 1825 to the present day the formation of unions has been legal in England. However, during the half century immediately following the repeal of the Combination Acts, the activities of each union were narrowly circumscribed by common law doctrines, which condemned as criminal conspiracies all agreements to raise wages and prices or to restrain trade. Since the unions were formed for the express purpose of winning better terms of employment, and that by collective agreement and collective action, their members were continually falling foul of this doctrine of conspiracy. Nevertheless the union movement spread with great rapidity. During the thirties it was caught up in the widespread social agitation of that decade and wasted its energies in support of schemes of social and political revolution. The collapse of the Chartist movement had a sobering effect upon labor leaders. The labor movement turned its attention to the problems of perfecting its organization and improving economic conditions within the different trades. Isolated unions embracing similar grades of labor in different sections of the country began to coalesce into nation-wide organizations. As these unions grew in strength however their normal policies of collective bargaining and their weapons of warfare — the strike and the picket — brought them more and more into conflict with the courts, which were prone

to view all such group activities as evidence of conspiracy to harass the employer and to interrupt the free play of competitive forces in the labor market. Largely in response to this common problem, the various unions of the country met in 1868 in a Trade Union Congress to establish a comprehensive working agreement on a national scale. This first Congress was attended by representatives of 118,000 unionists; in the Congress of the following year 250,000 unionists were represented; in 1872 the number had grown to 375,000. Since 1872 the annual Trade Union Congress has been an established feature of the labor movement of Great Britain.

Recent developments in Great Britain. The period from 1870 to the present day has been marked by several lines of development in the British labor movement. In the first place there has been a steady growth in union membership and a continuous improvement of union structure in the direction of integrating the various groups into strong national organizations. By 1920 the labor movement had attained a membership of approximately eight and a quarter millions. The importance of this movement as a factor in British national life is indicated by the fact that it embraced in this year sixty-five per cent of all wage-earners engaged in industrial occupations and over eighty per cent of the male workers in these occupations. In the most important branches of industry — mining, transportation, the mechanical trades, and the building trades — the workers are united in great nation-wide organizations, each of which includes virtually all the available labor in its particular department of industrial activity.

In the second place the law has been remodelled along lines more and more favorable to trade union policy. A decisive step was taken by the government in 1875, when the conspiracy doctrine was modified to exempt the unions from all penalties for peaceful collective action. The unions, as corporate organizations, were no longer punishable for forming agreements to raise wages by means of their power to control the supply of labor nor for engaging in strikes in furtherance of this aim. When, in the early years of the twentieth century, the courts showed a disposition to weaken these legal guarantees by interpretations of the law adverse to

union activity, the matter was set finally at rest by the Trade Disputes Act of 1906. By the terms of this Act the courts are debarred from entertaining suits against trade unions, as such, for any cause whatsoever. The individual members, like all other citizens, are punishable under the criminal law for violence to persons and property, but in their collective capacity they are immune from legal interference.

In the third place the British labor movement has become a powerful factor in the political life of the nation. Following a fitful and sporadic attempt to found an independent labor party in the nineties, organized labor entered the political arena during the first decade of the twentieth century in a determined manner. During the past quarter century the Labor Party has grown in power with greater rapidity than any other party in the political history of modern democratic nations. The momentum of this movement is shown in the following table.

Year	*Representatives in Parliament*	*Total Popular Vote*
1906	29	323,000
1910	42	505,000
1918	70	2,244,000
1922	144	4,000,000
1924	192	4,500,000
1925	151	5,500,000

In a three-sided election early in 1924 the Labor Party, though it commanded a minority of the votes of the country and a minority of the seats in Parliament, was called upon, as the most powerful single political group, to form the ministry and conduct the affairs of the nation. This position of precarious dominance was lost in the election in November of 1924, but the party increased its poll of votes and was recognized as the official Opposition of the Conservatives, who were then in power. In the political branch of its activities the British labor movement is avowedly socialistic, relying however upon peaceful and constitutional means to effect its ultimate aim of thorough social reconstruction.

Early history in the United States. The labor movement in the United States was both later in origin and of slower growth

than in Great Britain. The appearance of the factory system, which supplies the conditions indispensable to the formation of labor organizations, was delayed by the predominantly agricultural character of economic life and by the institutions of the frontier, which prevented the formation of a stable wage-earning class. No unions of the modern type existed during the colonial period of our history. Shortly after the adoption of the Constitution the industrial life of some of the larger cities began to show signs of social division between wage-earner and employer and the beginnings of group action among the former. Small and ephemeral unions developed among the skilled craftsmen of New York, Philadelphia, and Baltimore which, having for their primary object the maintenance of benefits for the sick, devoted a part of their effort to improving their working conditions. These sporadic movements were a feature of city life throughout the first quarter of the nineteenth century, but they resulted in no permanent labor organizations, and in no case did the union grow to sufficient size to become an influential factor in industry. Between 1825 and 1850 a more vigorous trade union movement arose, with its centre in those cities where the wage-earning class was rapidly increasing. By 1836 one hundred and sixty unions, with a combined membership of 300,000, were in existence in the seaboard cities. Trade associations were formed within the cities, uniting the individual craft groups for common action, and attempts were made to bring together the geographical units of each trade into a national grouping. The movement was drawn into politics, both within the cities and on a national scale; it also was misled into supporting many of the utopian and communistic experiments which feature this period of our national life. Weakened by its failure in these political and socialistic ventures and by the disastrous panic of 1837, all of the national and city-wide associations and most of the small local craft unions passed out of existence before 1850.

Adaptation of union structure to early economic conditions. Before the middle of the nineteenth century then no permanent organizations of labor had made their appearance in the United

States. After 1850 however the quickened pace of industrial development produced a labor movement which, despite decline in strength during years of panic and industrial depression, has continued to grow down to the present day. Like their predecessors, these unions were at first small combinations of craftsmen within a given trade and locality. Called into existence as experimental adjustments to the life problems of the workers and not in response to abstract notions of social organization, their forms of structure were adapted to the prevailing economic conditions of the country. At this time the typical industry was a small scale unit whose influence did not reach much beyond the limits of its own town or city. Moreover the working class was still divided for the most part into non-competing skilled groups, each performing a distinctive function and having few interests in common with other and different occupational groups, even within the limits of the locality. Within these narrow ranges the workers were thrown into competition with each other in their search for employment and thus tended to develop feelings of common interest in opposition to the interests of the local employers. The small, local, narrowly exclusive organizations of skilled craftsmen, which appeared to be the only type of union capable of surviving at this time, owed their vitality to the fact that they represented a better adjustment to the economic conditions of the age than did any broader or more comprehensive grouping.

Nation-wide organizations. However the social forces working within the country were destined soon to stimulate an organization of labor on a larger scale. Steam transportation broke down the geographical barriers which had divided the labor market into small, self-contained units and broadened the area within which craftsmen were brought into competition with each other. In response to this development, the local unions began to combine along craft lines into national organizations, each limiting its membership to workers in a specific trade or calling. By 1870 upwards of thirty skilled trades had succeeded in integrating their local units into unified organizations covering the entire national area. More powerful forces tended to hasten this line of develop-

ment. Beginning about 1870, business enterprise entered upon that process of nation-wide integration which is so prominent a feature of the modern industrial world. Gigantic industries, formed by the consolidation of formerly competing units or by the expansion of an individual enterprise, grew to positions of dominance in many lines of productive activity. The workers, scattered far and wide over the country, were brought face to face either with a single employer or with a number of employers so small that they could easily form working arrangements governing the terms at which they offered employment to the wage earners. Problems of wages, hours, and working conditions could no longer be viewed by each local group of workers as problems applying to themselves alone; the area of common interest broadened to conform to the increased scale of the typical business enterprise. Stimulated by these changes in the underlying economic organization of the country, the tendency toward national organization within each trade or craft proceeded at a rapid pace. At the present time the nation-wide union of workers in each separate trade is the standard form of labor organization; in 1920 there were 112 such national unions in existence.

Comprehensive organization of the wage class. But the forces which we have sketched were tending toward a still more comprehensive grouping of the workers. With the growth of national unity and the widening range of government activity in the sphere of industry, many problems made their appearance which touched the welfare of the wage-earners in many different trades simultaneously. Consequently it became desirable for these different groups of workers to perfect some sort of organization which would bring about coöperative action in the face of these common problems. Reacting to these forces, the labor movement in the United States from 1870 down to the present day has striven to set up a comprehensive, class-wide organization with sufficient coherence to weld together the entire mass of wage-earners. At first — during the seventies and eighties — the experiment was tried of discarding all lines of craft division and uniting the working people into a single centralized organization. For a time this movement, as measured in

terms of numerical strength, was successful. In 1886 the Knights of Labor, as the centralized union was called, numbered 600,000 members. But the grouping proved too comprehensive to produce coherent and harmonious action. Driven by reason of the heterogeneous character of its membership — skilled and unskilled, men and women, farmers, wage-earners, and professional men were enrolled together on a plane of equality — to support broad social and political policies, the Knights of Labor wasted its energies in sympathetic strikes and in ventures in the field of coöperative production which ended disastrously. Internal dissension arose within its ranks, and the organization began to disintegrate before 1890.

Its decline in power was accompanied by the rise of a somewhat different type of comprehensive union organization — the American Federation of Labor — which today dominates the labor movement in the United States. As its name implies, this organization is a federation, not a centralized national structure. It recognizes the necessity of separate national unions in each of the occupational groups. These autonomous craft unions have command over their own policies and programs of action, but they unite with each other to form a loose federal organization, for the purpose of solving problems which they share in common and of promoting unity of action within the ranks of the organized labor class as a whole. This partial merging of the whole labor movement into a single organization, though it lacks the features of close integration which were characteristics of the Knights of Labor, is nevertheless an indication that the trend of social forces in modern society is toward the creation of a distinct class of wage-earners with a consciousness of common interests.

Types of union structure: the trade union. In point of structure, union organization is an adjustment to the prevailing economic conditions of the time. Since these underlying conditions do not long remain stable in our rapidly changing world and since they are never quite the same in all branches of industry within a country or in different countries at a given time, the structure of the labor movement is constantly changing with the passing years and

is never uniform throughout the whole wage-earning population at any chosen date. It is possible however to distinguish two types of labor organization standing somewhat in contrast with each other, which have been illustrated in the preceding sketch of the historical development of the labor movement.

The first of these is the *craft* or *trade union*. This type attempts to bring together all the workers who perform the same specific function and who are therefore likely to compete with each other. For example, bricklayers, carpenters, plasterers, lathers, painters, though they all belong to the same major branch of industry, the building trades, are organized into separate unions in each locality. Similar craft divisions are to be found among the workers on the railroads. The locomotive engineers, the firemen, the trainmen, and the conductors are united into self-contained groups distinct from each other. In fact the craft union is firmly established as a type of organization in all branches of industry where skilled craftsmanship has survived the effects of machine technique. The local subdivisions within each craft have then united into national organizations in response to forces tending to broaden the range of the labor market. Finally, in order to bring about coöperation between the different unions, reliance is placed in the device of *federation*. In each industrial centre the local craft unions are federated into a city council for the purpose of solving problems of common interest in the locality. Similar federations within the area of each state attempt to bring about coherent action on legislative programs which fall within the power of the state legislature. Finally the American Federation of Labor draws the whole system together into a national grouping whose chief component units are the national craft unions.

The industrial union. The second major type of union structure is the *industrial union*. This form attempts to obliterate craft and trade divisions among the workers and to make membership in the union coterminous with the limits of the various industries. From what the reader now knows of the economic organization of the modern world, he will understand how the factory system, with its machine technique and its refined division of labor,

also, international union.

has tended to break down craftsmanship and undermine the position of the skilled worker. This process has been accompanied, in the United States, by a mounting tide of immigration, which has flooded the labor market with alien unskilled workers, whose traditional standard of living reduces the plane of competition to a level destructive of the higher standards of the native born artisans. When skilled and unskilled, native and alien, male and female struggle together to obtain employment, older lines of division between occupational groups tend to disappear, and the whole group of wage-earners in each major line of industry tends to develop a consciousness of common interests. Organization along craft lines is out of place in this setting; adjustment to the conditions of the workers' life calls for a type of organization which will bring the workers in each branch of industry into a single union regardless of differences in the specific functions performed.

A prominent example is to be found in the mining industry. The United Mine Workers of America, in its local branches, brings together all the workers who have any connection with the production of coal. Stationary engineers, carpenters, draymen, unskilled laborers, as well as the actual miners of coal, are united in a single comprehensive organization. So it is in the garment making trades, where the Amalgamated Clothing Workers represent a single union of wage-earners engaged in all the different processes which are involved in the production of clothing. The industrial union, like the craft union, consolidates its local units into national organizations, which attempt to perfect a monopoly of labor in each branch of industry throughout the country as a whole. Some of these national industrial unions have joined the American Federation of Labor in the desire to coöperate with the national craft unions in the solution of problems affecting the whole class of wage-earners; but their presence in this organization is an exception to the prevailing craft unionism upon which the Federation is founded.

Whether the trade union or the industrial union will survive in the future as the dominant type will depend upon which one

affords the more useful adjustment to the conditions of the industrial world.

Political activity of American unions. In one regard the American labor movement has not followed the lines of development marked out by the history of the older movements of European countries. During the past three decades, in which period the unions in the United States have made their most substantial and continuous gains in strength, they have abstained from organizing an independent labor party. Most of the larger national craft unions, as well as the American Federation of Labor, have incorporated in their constitutions provisions which forbid their members from attempting to resolve the organization into a political party. The American unions have conceived their problem to be primarily economic rather than political; their objective has been to enhance their bargaining power as against the employers and not to capture the power of the state in the effort to put through far-reaching reforms. Accordingly their members are urged to retain their independence of choice in regard to party membership, the official action of the unions in the way of influencing this choice consisting only in their endorsement of those candidates for public office presented by the major parties who have shown themselves "friendly to the interests of labor." Hence there is no distinctive American labor party comparable to the Labour Party of Great Britain.

Obstacles to political action. This difference is to be explained in part by reference to the past experience of organized labor in this country and in part by peculiarities in our governmental system. Several of the early attempts to organize labor in this country were destroyed by ill-advised campaigns in politics, and the unions have learned to fear the divisive influence of partisan politics upon their membership. On the other hand our governmental structure is such as to discourage independent activity on the part of any economic group. The division of sovereignty between national and state governments, with the important power over social legislation distributed among the forty-eight different commonwealths; the separation of the governmental machine into

semi-independent executive, legislative, and judicial units; the power of a Supreme Court, remote from political influences, to nullify social policies of the states and federal lawmakers on the ground of unconstitutionality; these and other characteristics of our system multiply the obstacles in the way of effective political action on the part of any minority economic group operating independently. Moreover the predominately *laissez faire* character of our governmental policy has tended to minimize the importance of legislation as a means toward the accomplishment of economic ends and has encouraged workers and employers alike to rely upon their own bargaining power.

Possibility of future political activity. However the labor movement in the United States does not conform to uniform pattern with regard to this question of political activity any more than in other matters of structure and policy. A few of the strong unions have consistently asserted the necessity of carrying on independent political action along lines similar to those followed by the labor movement of England. Since 1920 several attempts have been made to break down the opposition of the more conservative unions to this line of policy, and definitive labor parties have entered the political arena of the nation and of the various states. On the whole these efforts have proved ineffectual, especially when carried out on a national scale. It is noteworthy that the unions which sponsor the formation of an independent labor party in this country are almost wholly unions of the industrial type, whose broad inclusive basis of membership compels them to adopt policies of general social import. The tendencies which further this form of unionism therefore will probably promote the political activity of organized labor.

Other tendencies may be observed working in the same direction. In all industrial countries governments are becoming an increasingly important force in economic affairs, regulating business activities, establishing standards of economic conduct, restricting individual initiative along many lines, entering the industrial world as owner of productive enterprises and as employer of labor. These expanding economic functions of government

may well attain such a degree of development that organized labor can no longer accomplish its aims by economic action unsupported by its organized political power. In the field of politics, as in the industrial sphere, the structure of the labor movement will be worked out by the process of adjustment to the changing problems of the workers. If the independent labor party proves in the course of time to be a serviceable instrument for solving these problems, we may expect the unions to adjust their structure to meet these conditions.

Extent of the American labor movement. The following table shows the growth of unionism in this country since 1900.[1]

Year	*Total Membership of American Unions*
1900	868,500
1901	1,124,700
1902	1,375,900
1903	1,913,900
1904	2,072,700
1905	2,022,300
1906	1,958,700
1907	2,122,800
1908	2,130,600
1909	2,047,400
1910	2,184,200
1911	2,382,800
1912	2,483,500
1913	2,753,400
1914	2,716,900
1915	2,607,700
1916	2,808,000
1917	3,104,600
1918	3,508,400
1919	4,169,100
1920	5,110,800
1921	4,815,000
1922	4,059,400
1923	3,780,000

As will be seen, there was a growth in membership of more than three hundred per cent during the period covered by the table. The maximum increase occurred during the war years, especially

[1] Taken from *Growth of American Trade Unions*, by Leo Wolman, National Bureau of Economic Research.

at that time when the United States was engaged in war or was feeling the effects of the shortage of labor immediately following the war. The movement as a whole reached its high water mark in 1920, when it included upwards of five million members. After that a decline set in, which continued during the following three years. A careful examination of the table discloses a close correspondence between the strength of the union movement and the industrial activity of the country. Periods of general business depression witness a falling off of trade union membership — for example, 1904–1906, 1908–1909, 1913–1915, and most noticeably, 1920–1923. A study of the earlier history of the union movement in this country and elsewhere will show the same correlation. Thus in 1837, 1857, and again in 1873, crises destroyed an American union movement which had flourished during the preceding years. In each case the loss of membership was more than made up during subsequent " prosperity " periods.

Union membership and the business cycle. The reason for the correlation between the strength of the union movement and the state of industrial activity is not far to seek. The whole movement, as we have seen, must be viewed as an adjustment device intended to better the conditions of the wage-earners. Its appeal to the workers depends upon its success in achieving tangible results in the form of higher wages, shorter hours, or greater economic security. When the underlying economic trends in the industrial structure of the country as a whole are such as to strengthen the economic position of the wage-earners, the organized labor movement gets the credit for advances in wages and other improvements which are in reality the product of these basic forces. During the prosperity period of the business cycle, when the demand for labor increases, when employment becomes more stable and continuous and wages rise, the workers are in a strategic position to press their demands upon the employers, and their efforts meet with unusual success. The unions are enabled to demonstrate their utility to the rank and file in such practical terms that their membership grows with great rapidity. On the other hand, when the reaction sets in and the trend of economic

forces undermines the bargaining power of the laboring population, the unions wage a losing battle against falling wages, instability of employment, and other reductions in economic welfare. Those elements of their membership which are at all times lukewarm in their loyalty desert the unions during these periods of hard times, and many others are forced to withdraw because of their inability to continue payment of union dues. Thus the strength of the movement fluctuates with the ebb and flow of the business cycle. But it is noteworthy that the unions complete each cycle with a net gain in strength, and that, despite cyclical variations, the movement grows in strength steadily and consistently over long run periods.

Relative importance of organized labor. More reliable as an indication of the importance of the organized labor movement than statistics of union membership is the proportion which this membership bears to the whole industrial population of the country. Analysis of the census of 1910 shows that the industrial wage-earners numbered in that year about 19,000,000; by 1920 this number had increased to approximately 23,500,000. At the earlier of these dates about 11 per cent of the whole wage-earning population were members of the unions; by 1920 about 21 per cent had joined the movement.

The organized labor movement obviously represents a minority of the working people of the country. But its relative strength is much greater than one might infer from these percentages. The great body of wage-earners contains many groups of workers whose conditions of life do not promote collective action and many other groups attached to subordinate or small scale enterprises whose activities, even when they are organized, exert but little effect upon the functioning of the industrial structure as a whole. The unions have attained greatest strength among the workers in those "key" industries upon whose smooth functioning the economic welfare of our society rests, a fact which greatly enhances the power of unions. This fact is disclosed by the following table which records the percentage of union membership in certain selected lines of occupation.

PER CENT OF WAGE-EARNERS ORGANIZED IN SELECTED INDUSTRIES

Industry	*Per Cent Organized* *1910*	*1920*
Coal Mining	35.0	51.0
Steam Transportation	23.5	57.5
Water Transportation	28.9	85.5
Electrical and Street Railways	21.8	52.9
Clothing	16.9	57.8
Building Trades	16.4	25.5
Iron and Steel Manufacture	10.4	28.1

Even this does not tell the full story, since many of these major divisions of industry contain occupational subdivisions whose labor is essential to the operation of the industry and whose percentage of union membership is higher than that for the division as a whole. Thus in steam transportation 57 per cent of the workers were organized in 1920, but of the locomotive engineers and firemen over 75 per cent were members of the unions, while the railroad conductors were almost completely organized. Similarly in the building trades, in which one fourth of all the workers were organized in 1920, about half of all the carpenters, masons, and plasterers — skilled crafts whose labor is essential to the functioning of this industrial group — were members of unions. In these major divisions of industry the collective strength of the organized labor movement is strong enough to influence profoundly the economic activity of the country as a whole.

Where unionism does not flourish: among women workers. Among two groups of the laboring population in this country unionism has had slow growth. The first comprises the women workers. Until the last decade the trades consisting predominantly of women wage-earners were wholly unorganized, and the female membership of the unions in the mixed trades was negligible. In 1910 the total female membership of the unions in this country was but 77,000, or one and a half per cent of the number of women wage-earners. During the following decade organization among the women workers increased with some rapidity, the total in 1920 amounting to approximately 400,000, or six and a half per cent of the female workers of the country. Some of the causes of this

slow growth inhere in the psychology of the women, who on the whole are adverse to taking part in a distinctively working-class movement. The forces of custom and tradition developed in the homes from which the women workers come forbid them to commit themselves openly and avowedly to the life of the wage-earner and to profess loyalty to his group. Then too the women workers in this country as a rule view their life in industry as a temporary expedient occupying their time before marriage, and few have sufficient interest in their occupations to bear the financial burdens entailed by union membership. In the mixed trades the jealousy and suspicion of the male workers, who have been loath to accept women into membership in their organizations, is a factor in the situation.

The chief explanation of the slow growth of unionism among women workers however is not to be found in the factors we have mentioned but in the nature of the occupations in which women are principally engaged. Reference to the table on page 522 will show that the women workers are employed chiefly in professional, domestic, and personal service and in clerical occupations. These are pursuits where the effects of the factory system which stimulate collective action among the workers are least felt; even among the male workers unionism has not gained a foothold here. In the distinctively industrial occupations — such as manufacturing and transportation — over eighteen per cent of the women workers are organized, and the extent of organization is increasing steadily. It may safely be predicted that, if the tendency toward the increase of female labor in these occupations continues, trade unionism among the women workers will not long be delayed. In England, where this tendency has been more marked and of longer duration, unions of women workers numbered over four million members in 1920.

Among the unskilled workers. The second section of the working population which is on the whole unorganized in America consists of the unskilled workers. Craft unionism, the dominant type of labor organization in the United States, is limited of course to the skilled groups — the aristocrats of the labor world. There

is no natural place in the philosophy of this movement for the men without a specific trade. The situation is complicated by the fact that our unskilled workers are to a large extent aliens, men of foreign tongue and outlandish standards of life, who are viewed with contempt or antagonism by the Americanized artisans whose standards are higher. The newer industrial unionism finds a place for this lowest grade of labor in the same organizations with all other workers employed in each major branch of industrial enterprise, but this type of unionism is still relatively weak in our country. It should be noted that no consistent effort has been made to organize this large group of the lowest paid and most necessitous of the wage-earning population and to bring within their reach whatever economic benefits result from increased bargaining power. On the contrary, as we shall see in due course, many of the standard trade union policies are calculated to improve the condition of the skilled workers at the expense of this mass of the unskilled.

Non-economic activities of the unions. We are now prepared to study the activities of organized labor and to appraise its importance as one of the economic forces of our social order. These activities are many and varied. They are not all economic in character and hence do not all fall within the scope of our present study. No permanent association can be formed among men which does not tend to take on functions apart from its major purposes; for these associations will be made the mechanism through which their members give effect to the many social impulses which are common to mankind — the desire for companionship, for organized sport and amusement, for self-expression along many lines. As older community ties are broken down by the progress of industry, the labor union becomes, for many of the wage-earners, one of the few remaining bonds uniting him to his fellows in durable association, and it is made to serve purposes which among other groups of the population are served by clubs and fraternal organizations. These incidental functions of the unions are important and should be taken into account in any attempt to understand the full significance of this movement in

our social world, but we must pass them over here with no more than bare mention.

Benefit systems of the unions. One of the oldest of the economic functions performed by labor unions has been that of providing the wage-earners with a form of mutual insurance against some of the hazards of their lives. This function is discharged by what is called the "benefit system" of the unions. Most of the labor organizations of England and the United States have collected funds out of which to pay small incomes to their members in case of accident, sickness, or unemployment and to provide for burial expenses. At the outset these benefits were financed by a crude system of assessment which called upon the members for contributions as need arose. In all the older and more firmly established unions, however, the system has now been placed on an actuarial basis similar to that which underlies all scientific forms of insurance. The members make monthly contributions into the treasuries of their organizations, the resulting funds being segregated from funds held for other purposes and kept invested in safe securities for the purpose of financing the benefits to which the contributors are entitled. The unions vary in the amount of emphasis placed upon this phase of their service to their membership. Some do not have the benefit system at all, and others discharge it in a desultory and unreliable fashion. But among the more powerful national unions it holds a prominent place and is responsible, in no small measure, for the success which these unions have had in building up and retaining their membership.

As examples of the importance of the insurance principle in the labor movement, the unions among the workers on the steam railways may be cited. The Brotherhood of Locomotive Engineers carries approximately two hundred millions of dollars' worth of insurance in force at a given time, pays out some two and a half million dollars in benefits each year, and has made total payments to its members since the system was instituted amounting to over fifty million dollars. The Firemen's union has paid a total of twenty-seven million dollars in benefits and now carries about one hundred and fifty million dollars' worth of insurance continuously.

The Trainmen's union has paid a total of fifty-five million dollars since the benefit system went into effect. Other craft unions perform this function on a similarly extensive scale. In 1921 the national craft unions belonging to the American Federation of Labor — these do not include the unions just mentioned, since the railroad unions have not joined the Federation — paid out in benefits approximately five and a half million dollars.

The examples given will suffice to show the importance of this phase of union activity and to give some indication of the magnitude of the financial resource which the unions can wield by mobilizing the small individual contributions of their members. But important as is the function of mutual insurance in the lives of the workers, other organizations exist which discharge it quite as effectively as do the labor unions. Except in the case of the more hazardous occupations, insurance companies established as business enterprises offer the wage-earners indemnity against industrial accident and disease. It is true that in the United States there is, outside the labor unions, no organized effort on a wide scale to provide insurance against unemployment, but the experience of England has shown that this can be done through other mechanism than that provided by the unions. The unique and distinctive economic functions of the labor organizations therefore do not inhere in their benefit systems. They consist in the attempts of the unions to raise wages, shorten hours, and increase the economic security of the workers through the use of collective bargaining.

Collective bargaining. This device of organized labor has been defined and explained in connection with our study of the law of wages. The general argument in its favor, as compared with the bargaining of individual workers, has also been presented, for we have had reason to conclude that the isolated worker is at a disadvantage in the wage bargain with an employer and that this inequality in bargaining power is likely to result in a wage rate lower than the economic worth of the worker. We may pause here only to note that this natural superiority of the employer is, in special circumstances, increased by conditions peculiar to the

locality or the enterprise. Such is the case for example when the workers are of foreign birth, ignorant of the language and industrial conditions of the country and habituated to a low standard of living, or when they are women or children. Similar effects follow when the employer, in addition to his ability to supply employment, has other holds on the worker which he may use as weapons of intimidation, such for example as his ownership of all the available homes in the village or his position as creditor for arrears of payment due at the company store. We should note furthermore that the forces which make for the normal bargaining weakness of the worker grow in importance with the growth of the factory system. The larger the scale of the enterprise, the less competent is any individual worker to defend his interests.

Gains through collective bargaining. Our study of the law of wages has made it clear that the workers may raise their real wages by a change in bargaining methods which substitutes the collective for the individual wage contract. It has also marked out the limits within which this improvement can occur. Wages can be raised by collective bargaining only if they were previously lower than the marginal economic worth of the workers; they can be advanced beyond the point of marginal productivity only at the cost of throwing some of the laborers out of work.

Collective bargaining however is not of advantage to the laborers alone; it results in certain indirect benefits to the employer, which enhance its value in the industrial system. The mechanism through which this type of bargaining is carried out may be turned to account in drawing workers and employers together into closer contact on a basis of mutual tolerance and good feeling. Many of the wastes in modern industry and much of the friction and bitterness engendered in the field of industrial relations flow from the absence of such contact. Organization of the workers for collective bargaining and the establishment of a permanent structure for negotiation with the employers may effect a real improvement in the situation. The workers are supplied with a mechanism for voicing their grievances and are assured of a respectful hearing.

Discipline in the shop is more effective when it bears on observance of the terms of a contract in the making of which the men have had equal voice. The subordinate officials are restrained from harsh or overbearing tactics; fair and tolerant treatment of the workers becomes good policy in view of their organized power. Above all labor is given a recognized share in the control of those features of industrial policy which most vitally affect their welfare, and this cannot but influence their attitude toward their jobs and toward the enterprise. It goes far toward supplying a substitute for that personal interest in the product and the fortunes of the concern which existed in the handicraft stage of economic evolution but was destroyed, with far-reaching results, by the growth of the factory system. These benefits from organized collective bargaining are now recognized by practically all modern employers whose enterprises have grown to any size. In England and Europe and to an increasing extent in the United States as well the large employer who insists on dealing with his workers individually is recognized as hopelessly behind the times.

Conflict over the method of collective bargaining. But though the soundness of the principle of collective bargaining is generally conceded, there is no such agreement regarding the method of applying this principle. One distinctively American phase of the problem of industrial relations is the quarrel over the question of "recognizing the union." Organized labor insists that the workmen in each of the different enterprises must be represented in their bargaining by the regular officials of the union. These men are not on the payroll of any employer with whom they negotiate wage agreements but are professional bargainers representing all the members of their union within a given geographical area. Such a system of bargaining presupposes, in its completed form, organization of employers in nation-wide associations paralleling the national unions of workmen, and the negotiation of national agreements between representatives of the two organizations, binding upon the separate enterprises in their different localities. Recognition of the union implies the willingness of employers to adopt this method of bargaining.

This use of professional bargainers is defended on general principles as well as on the ground that it results in certain specific advantages to the workmen. Acceptance of the principle of collective bargaining, it is asserted, implies the right of the workers to be represented by men of their own choosing; the benefits of the system will be nullified in practice if this right of choice is restricted in any way.

On the other hand there are certain employers and powerful employers' associations in the United States which refuse to accept this method of applying the principle of collective bargaining. They insist on dealing only with their workmen through representatives of the latter who are at the same time employees in the enterprise. The business agents, or "walking delegates," of the union are opposed on the grounds that they are ignorant of the conditions of the specific enterprise and are aggressive and domineering to a degree which makes peaceful compromise impossible.

When the question is considered on its merits, divorced from other aspects of union policy, the case in favor of the union method of collective bargaining is conclusive. It is apparent, on the face of the matter, that the interests of the workmen cannot safely be entrusted to a representative who is dependent on the employer for his livelihood and who may be victimized by petty persecution of the foremen or discharged outright for diligence in the performance of his duty to the union. This method might in time easily result in the elimination of all representatives who were not sufficiently pliable to be acceptable to the employer, in which case the negotiations could not be called true bargaining. There is some inconsistency in the attitude of those employers who refuse to confer with union officials while yet acknowledging the validity of the collective bargaining principle. On their side they habitually carry on negotiations with labor through the service of managers who are employed for this purpose, not because they are stockholders or members of the ownership faction of the enterprise, but because they are experts of professional status. When the workmen designate a union official as their representative they are imitating this established business practice.

However it is virtually impossible to consider this question solely on its merits, for the controversy over the method of collective bargaining touches the very life of the union itself and is colored by all sorts of considerations relating to the social utility of unionism. If the workmen were allowed to organize for collective bargaining only within the limits of each individual enterprise and were compelled to employ ordinary wage-earners as their representatives, the union movement — in the form of class-wide labor organization — would lose its reason for existence. Opposition to the union method of collective bargaining is most frequently inspired by antagonism to certain practices of organized labor — *e.g.* restriction of output and make-work policies — and has the ulterior motive of destroying or enervating the union movement. On the other hand the officialdom of the unions base their case in defense of this method of bargaining, not alone on its undoubted superiority over the alternative method, but also on broader grounds of the utility and social desirability of the labor union.

The closed shop. Before discussing the controversy which rages around the question of the closed shop, it is necessary for us to define our terms with some care. Strictly defined, a closed shop is one in which none but union men will be employed to perform those tasks whose labor supply has been unionized. Its opposite, the open shop, is one in which there is no discrimination in offering employment as between union and non-union men. Both of these terms are often ambiguously used. For example some employers who declare their opposition to the closed shop announce that they will not hire union men at all. This policy they call an " open shop " policy; whereas in reality it establishes a type of closed shop — one that is closed to union labor. Moreover the closed shop has been confused with the *closed union* in much discussion of its merits and defects. A closed union is one which refuses admission to properly qualified workers. The effects of the closed shop are obviously quite different when this privilege is enjoyed by a closed union than when the union is open to all comers. But it is a mistake to assume that the closed shop and closed union are necessarily found conjoined.

The closed shop is merely an extension of the union program of collective bargaining with its implication, standardization. Collective bargaining is based on the assumption that the resulting agreement will be binding on the entire labor force. If it applied to only a fraction of the workers in an enterprise, the established standards might be disregarded by the employer in his dealings with the non-union workers. When this resulted in lower wages for some men than for others, it is obvious that preference would be given to the former whenever the two groups — non-union and union workers — were competing for employment. Hence the purposes of collective bargaining and the effectiveness of the union would be thwarted. From the point of view of the corporate interests of the union, there are obvious advantages in the closed shop policy. Clearly it sets at rest the question of recognition. It enhances the worker's inducement to join the union and reduces the difficulty of expanding the field covered by organized labor. It strengthens the disciplinary powers of the union officials; the recalcitrant member expelled from the union suffers unemployment. This in turn facilitates the collection of union dues and makes for faithful observance of union rules. It is an absolute guarantee against the victimization of spokesmen for the workers when defending the interests of their constituents. It strengthens the hand of the union when, in the last resort, it orders a strike. In fact the closed shop makes the strike so formidable that the threat of the strike often suffices to win the point for which the union has been contending.

Advocates of the closed shop policy urge in its favor certain advantages to the employer. First it makes for peace in the shop. Union and non-union men can rarely work side by side in harmony; the loyal member of the union looks upon the non-union worker as a renegade, a traitor to his class. Secondly it gives the union power to enforce faithful observance of the collective agreement and to guarantee a standard of performance. Thirdly it insures the good will of the buying public in the case of all those concerns whose products are sold chiefly among working classes; from this point of view the closed shop and the union label go hand in hand.

Finally it protects the " fair " employer against cutthroat competition from enterprises which mistreat their workers. This last argument holds good obviously only on the assumption that the closed shop applies to all units of a sharply competitive industry.

The economic arguments against the closed shop cohere about a common point; namely, that it places dangerous power in the hands of union officials and increases the menace of all forms of anti-social practice in which the unions are tempted to indulge. Much is made of the fact that the closed shop weakens the efforts of management to stimulate the efficiency of the working force. The employer's great disciplinary weapon is the right of discharge. It is asserted that the closed shop nullifies the effect of this weapon. Where this is true however it does not follow as a necessary consequence of the closed shop itself, but results from the indifference of the union to the efficiency of its members. If we could imagine that all unions were intent on supporting the efforts of management to reduce cost and save waste, it would be apparent that a closed shop policy would increase the disciplinary powers of the employer and act as a positive force on the side of efficiency. So it is with the other charges in the indictment of this policy; for example that it encourages the make-work practices of laborers, solidifies their opposition to labor-saving machinery, results in many monopolistic abuses. When we examine these charges, we find that they are directed against the uses to which the unions put their power, rather than against the specific device of the closed shop. Thus the final judgment on this phase of the union program must turn on the question of the general social philosophy of organized labor and the general economic behavior of unionized workers.

Restrictive tendencies of union policy. Thus far we have considered the economic program of organized labor merely as a collection of devices for increasing the bargaining power of the wage-earners. Collective bargaining and the closed shop are devices useful to labor in that they enable the workers to obtain the full reward for their services as set by economic law. They are all quite harmonious with an open union policy and in fact with the absence of any restrictive or monopolistic devices in the program

of organized labor. But unless supported by such devices, their effectiveness to improve the economic condition of labor is strictly limited. The real income of the different groups of workers is controlled fundamentally by the numbers in these groups in relation to the demand for their services. An overcrowded group may make full use of collective bargaining, standardization, and the closed shop and yet, if all its members are to find employment, must be content with a miserably small income. Recognizing this fact, the strongly entrenched unions are prone to adopt various monopolistic practices which are similar in social effect to those employed as a matter of course by business men who are in position to control the market supplies of their commodities.

These restrictive features of the union program fall into two major groups, each calculated to enrich the members of the union at the expense of some other class in society. In the first place some unions adopt the practice of artificially restricting their membership. This is done, not by a direct refusal to admit qualified workmen, but in certain indirect ways which reach the same result; for example by setting up prohibitive initiation fees or requiring an unreasonable period of apprenticeship. It should be clear that no union dares adopt such a policy until it has obtained a sufficient degree of competitive power. The closed union cannot exist under open shop conditions, for the reason that the members of the union would then be compelled to compete for employment in the same industries with the excluded workmen, whose competition would be the more destructive of union standards because they were unorganized. Also the right of collective bargaining through union officials must have been won before there can be any incentive to close the union; for obvious reasons the employers' method of applying the principle of collective bargaining is effective against the closed union. Collective bargaining and the closed shop, though innocuous in themselves, take on a sinister aspect when they are used as the indispensable first steps toward the creation of a monopoly of labor. By applying the policy of restricting its membership, an organized group may make of itself a short factor in industry, raising the value of its marginal

man's product and hence the wages of all its members. These gains are made at the expense, in the first instance, of the excluded workmen who are forced to render service of a type which is more plentiful and of lower value than what they are qualified to perform. In the long run all members of society share the loss through the decline in the social income which must result from employing a part of the labor supply at relatively unproductive tasks.

The second type of restrictive practices consists in a variety of devices for reducing the normal output of union members. These have been alluded to in a previous chapter. We have seen that all wage-earners are disposed, under certain conditions, to restrict their output and adopt "make work" policies. In large part this is a reaction against the menace of unemployment; in part also it is a deliberate attempt to enhance the value of the product by increasing its scarcity. Restriction of output therefore is not a unique policy of organized labor; though it is true that the trade union is in position to apply the policy with greater effect. It must not be inferred that all unions practice the two forms of restriction which we have discussed. Some unions renounce their use on grounds of broad social welfare; others have not attained to sufficient strength to make them effective.

The strike. Conflict in the field of industrial relations is forced upon the attention of the public when the bargaining system breaks down and employers and workers attempt to settle the issue by open warfare. Each of these groups holds a weapon in reserve to be used in the last resort; the employers possess the *lockout*, the workers, the *strike*. These are correlative weapons. The lockout occurs when an employer, or employers' association, withdraws employment from a group of workers pending the settlement of some issue which is in dispute. The strike is a collective withdrawal of labor by a group of workers until the employers consent to certain demands which have been made upon them. These weapons are alike in that their use has the effect of stopping production and throwing a burden of loss upon the public. But since the lockout is less frequently used than the strike, it does not receive the same amount of public attention.

Spontaneous quitting of work is the last resort of wage-earners when all other methods fail to win for them acceptable terms of employment. The unions have carried this weapon over into their own program of action, regularized its use, and made it more effective. All strikes involve loss to the workers; hence the weapon is not used as a matter of course but in general only after deliberation and in defense of what the workers conceive to be a vital element in their standard. It remains true nevertheless that the menace of the strike has increased in direct ratio with the spread of unionism. Strikes occur much more frequently today than when the union movement was in its infancy; they take place on a wider scale and are more protracted and more bitterly waged; compromise through arbitration and other peaceful methods of settlement tends to be scorned by the workers as they perfect their organizations. Organized labor, like the nation, prepares for war in times of peace, collecting reserve funds with which to finance the strike, perfecting the machinery for carrying on the conflict, and educating their members in ideas of discipline and mass action. They decide to strike, not in desperation, but after a cool calculation of the chances of success — a calculation in which loss to the public does not enter. As unionism spreads, each strike becomes something of a national calamity, embracing a wider section of the industrial area and paralyzing the sources of supply for goods essential to public health and safety.

Statistics of strikes. The magnitude of the strike problem will be made clear by a glance at the following table.[1]

TREND OF STRIKES IN THE UNITED STATES, 1881–1921

	Average Per Year		*Relative Numbers (Average 1881–1885 = 100)*	
Years	*Strikes*	*Strikers*	*Strikes*	*Strikers*
1881–1885	498	124,005	100	100
1886–1890	1,336	255,863	268	206
1891–1895	1,337	278,868	268	225
1895–1900	1,351	406,639	271	327
1901–1905	2,793	1,547,087	561	1,324
1915–1921	3,043	1,744,502	611	1,407

[1] Adapted from an article by Paul H. Douglas, "The Trend of Strikes," *Journal of the American Statistical Association*, Sept. 1923.

It will be noted that there was an increase of 511 per cent in the number of strikes per year during the period embraced in the table and a fourteen-fold increase in the number of strikers in the last five-year period as compared with the first. Of course the entire period was one of rapid industrial expansion, when both the number of business units and the number of workers was on the increase. The figures should be corrected by an index showing the increase in the number of industrial workers. If this were done a part of the apparent increase in the number of strikes and strikers would disappear; nevertheless there would still be over twice as many strikes in proportion to the working population in 1915–1921 as in 1881–1885 and nearly five times as many strikers. The total number of days lost through strikes rose from a yearly average of approximately three millions during the period 1881–1885 to over 50,000,000 yearly in the period 1915–1921. The growth of large scale industry, with the consequent broadening of the segment of the working population affected by strikes in a given enterprise, the development of strong national unions, and the unrest in certain basic industries are the principal causes of this increase in strikes.

Violence in strikes. The purpose of the strike does not demand that the workers resort to methods of violence. On the face of the matter the strike is comparable with the normal business practice of withdrawing a commodity from the market when its owner is dissatisfied with the terms resulting from bargaining. Yet each strike creates a situation in which violence to persons and property is very likely to develop, and many strikes have degenerated into a state of disorder closely bordering on civil war. It is to be noted that violence of this sort is rarely directed at the employer's person or his property; though property is often destroyed in the course of the disorder, this is usually an indirect result of the strikers' activity and not an intentional policy of the union. The chief causes of violence in strikes are two: the presence on the scene of conflict of non-union strike-breakers, and the interference of the public through its police force or militia. As pointed out in another place the free lance worker who accepts

employment in an industry during a strike is viewed as a "job-stealer," against whom methods of coercion may fairly be used when milder persuasive devices fail. As for the injection of public power into the dispute, in the form of police or militia, this is always viewed as a device of the employer to strengthen his own position. Appeals for public interference always come from the employers and property holders, never from the strikers and wage-earners. The appearance of the militia on the scene of action therefore is taken as evidence that force is to be used to aid the employer to win his case.

Though the number of strikes increases with the spread of unionism, the amount of violence in the average strike tends to decrease. This is due principally to two factors. In the first place as a union solidifies its position and expands its membership it has less to fear from the use of non-union labor in the rôle of strike-breaker. The strongest unions in the United States are virtually "scab-proof." The brotherhoods on the railroads are examples of this well-nigh impregnable position; the amount of non-union labor available to break a strike of the engineers, firemen, trainmen, and conductors could have no effect on the outcome of a conflict in this all-important branch of industry. Other unions, not so strong as these on a national scale, enjoy a virtual monopoly in many localities. Such are the building trades unions, the miners, the seamen and longshoremen. With the removal of the scab as a menace to the success of the strike, one cause of violence disappears.

The second factor in the spread of unionism which tends to diminish violence is a matter of the self-interest of the union. Violent strikes are rarely won by the workers. In the final analysis public sympathy is a determining factor in the outcome of strikes, and the public has no sympathy with violent methods. Knowing this and deciding on the strike policy after a careful calculation of the chances of success and failure, the strong unions take steps to suppress any violent or destructive tendencies of their members. Since union discipline varies directly with the degree to which they have perfected their organization, the expansion of the unions tends, for this reason as well as for the one just men-

tioned, toward a decrease of violence in the industrial conflict. The exceptions are supplied by those unions which have reason to feel that they are fighting for their right to exist because of repeated attacks of belligerent employers' associations, or by local units of a union whose members are bellicose and domineering above the average. On the whole the chief examples of violence today are to be found in the spontaneous strikes of unorganized and undisciplined workers who are fighting in desperation against the unusually burdensome conditions of their lives.

CHAPTER LIV

SOLUTIONS OF THE INDUSTRIAL CONFLICT

Protecting the public against strikes. So interdependent are the various parts of our industrial structure that no disturbance can occur in one field of economic activity without affecting other groups of producers and the consuming public. An abrupt cessation of the flow of goods from any basic industry — mining, transportation, the metal trades — may paralyze the whole productive system and, if prolonged, cause suffering and hardship resembling that consequent on civil or international war. Indeed there is no large branch of the industrial world which can be thrown into confusion without working injury to large numbers of people who are in no wise party to the original dispute. As trade unionism has spread, widening the area embraced in a single industrial conflict, this realization of the public interest in the relations of employer and worker has impelled the nations to experiment with various devices for the prevention or speedy settlement of industrial disputes. These devices may be grouped in an ascending order of increasing public intervention and coercion, with purely voluntary and informal arrangements at one end of the scale and the application of governmental compulsion at the other.

Failure of legal prohibition of strikes in Great Britain. Long experience has shown that attempts to prohibit strikes by law, if unaccompanied by constructive provisions for the settlement of industrial disputes, are ineffective. During the first quarter of the eighteenth century, when unionism was condemned in England as conspiracy, all forms of strikes were viewed as evidence of criminal restraint of trade and were rigorously suppressed. Even after the Acts of 1824–1825 had legalized the union, strikes were still condemned under the doctrine of conspiracy, and workmen engaging in them were punished. However the courts were unable to pre-

vent the workers from striking; the blind opposition of law had the effect merely of inciting the strikers to violence, thus aggravating the evil. Legal prohibition had to give way before a public opinion generated by those economic forces which made the strike an indispensable weapon for wage-earners under the factory system. With the removal of the conspiracy doctrine during the seventies, all forms of the strike have been legal in England except in certain public services vital to safety and health.

Legal status of the strike in America. In the United States there is no single law governing the activities of the unions, since this is a matter lying within the province of the states. As regards strikes, state laws are so frequently changed and the opinions of state courts are so at variance with each other that it is impossible to phrase a general rule regarding the legality of the strike in this country. The only approach to a general principle is this: that all strikes are legal, provided their object is clearly and solely to improve the economic condition of the strikers; *e.g.*, by raising their wages, shortening their hours, or bettering their conditions of work. To force an unwilling worker to labor, even to discharge a contract into which he has entered, is involuntary servitude, forbidden by the United States constitution. But this rule of law, though apparently clear on its face, does not legalize the primary form of strike in all circumstances and under all conditions. Indeed the state and federal courts have frequently attempted to solve the problem of the industrial conflict by the simple expedient of forbidding an open break between employer and organized employees, and they are still making this attempt. The courts of this country look to the motive of the strikers, not to their conduct, for justification of their action. Combinations are criminal conspiracies under our laws when their intent is to destroy property or to interfere with the rights of others. Since almost all strikes involve preconcerted planning and organization and must inevitably work injury to the employer and the public, it is not difficult for the court to find in them evidence of conspiracy dominated by malice. When the court thus assesses the motives, strikes are likely to be prohibited.

Failure of the injunction. The means now most frequently employed by our courts to prohibit strikes is the injunction process. The injunction is an order of a court restraining conduct which threatens irreparable injury to property. Violation of an injunction is "contempt of court," punishable by fine or imprisonment at the discretion of the judge and without jury trial. This device of law has had a long history in Anglo-Saxon jurisprudence, but it is used in no other country than the United States in connection with industrial disputes. It was first used in this way in 1888, and since 1900 the courts have used it with increasing frequency to prohibit union activities which were held to be of the nature of criminal conspiracies. In 1919 the injunction was used on a national scale in an effort to prevent a strike of the coal miners, and in 1922 a nation-wide strike on the railroads was enjoined by a Federal Court. In neither case was it effective. Indeed the experience of the United States, like that of England, has shown that no solution to the industrial conflict lies along the line of mere legal prohibition. Strikes have been broken by the use of the injunction, and in some instances the employer has succeeded, with the aid of the court, in winning his case against the opposition of the workers. But the causes of the evil have remained unaffected. The workers refuse and cannot be compelled to surrender this weapon of self-defense so long as the powers of the employer are left undiminished, or so long as no alternative machinery exists for the adjustment of grievances.

Conciliation. The least formal of the measures used to effect a speedy settlement of a trade dispute is *conciliation.* Strictly defined, conciliation denotes a method of settlement in which the employers and the workers, or representatives of these two groups, voluntarily meet together for discussion and agreement without aid or compulsion from the outside.

Conciliation implies a high degree of organization on both sides and a habit of collective negotiation. It is a natural outgrowth of the development of trade unionism on the one hand and employers' associations on the other. The method is best used for interpreting the meaning of doubtful points in an existing trade

agreement or of working out a peaceful compromise between conflicting demands of an employers' association and a trade union, to be embodied in a future agreement. Such differences as these usually relate to specific, practical matters of detail, on which compromise is both possible and beneficial to both parties. In the absence of machinery which begins almost automatically to function for the settlement of these differences, the alternative is an open break and a trial of force, which causes an immediate loss to all concerned and leaves an aftermath of bitterness and resentment. Questions of general principle, such as the right to organize, the prerogatives of union officials, etc., cannot possibly be settled by conciliation. Agreements to bring such matters within the range of conciliation almost invariably cause a breakdown of the whole machinery of friendly negotiation. Furthermore experience has proved that it is necessary for the conciliation board to reach a unanimous decision on all matters referred to it. The board is made up of equal representatives of workers and employers; its proceedings are informal; its decisions cannot be enforced against a recalcitrant party. Unless decisions are unanimous, the constituents of those representatives who have voiced dissent are very likely to repudiate the verdict. A strike or lockout ensues, which brings the whole process of conciliation into disrepute.

(A) **Conciliation promoted by government.** In many countries attempts have been made through governmental action to set up a permanent system of conciliation which will embrace all departments of the industrial field. Perhaps the best known of these is the Whitley Industrial Council Plan of Great Britain, so-called after the name of the chairman of the Parliamentary Committee which proposed the plan in 1917. The essence of the plan is recognition of the right of organized labor to an equal voice in the determination of those features of industry which most closely affect the welfare of the wage-earners. It provides for a hierarchy of standing joint councils, composed of equal representatives of the trade unions and employers' associations, in each major branch of industry. At the top of the structure is the national council,

which meets at regular intervals to discuss problems of general importance to the trade as a whole, such as standards of wages and hours, regularization of production and employment, collection of statistics and information. Regional or district committees, constructed along similar lines, consider problems peculiar to specific geographical areas. At the basis of the system are a series of joint "works" councils, one in each industrial unit, which provide machinery for continuous negotiation between local union leaders and representatives of management in the individual enterprises. The government has no official relation to the system beyond facilitating its adoption by voluntary action of the parties concerned. However, in harmony with the principle of unionism on which the whole scheme is founded, the government undertakes to encourage organization in those branches of industry where unionism has not obtained a foothold and to introduce temporary machinery for negotiation intended to develop into a full-grown system of joint councils when unionism has grown to full proportions. The system has been adopted by many branches of industry in Great Britain, but on the whole its adoption has been limited to the smaller and less important trades. The great national unions have preferred their own machinery for negotiation, while many of the stronger employers' associations have feared to adopt a system which implies a more complete recognition of the powers of trade unionism than they are willing to concede.

A somewhat similar system was introduced into Germany by law in 1919. This law requires the election of joint works councils in all establishments employing twenty or more persons, larger regional councils, and a national council in all industries which cover a wide geographical area. The German system is mandatory, while that of Great Britain is optional. The powers of organized labor in Germany were further extended by a law of 1921, which gives the workers the right to demand information regarding the financial policies of the business and the right to representation on the board of directors.

In the United States the use of conciliation has been retarded by the refusal of many employers to concede to the workers the

right to organize and to negotiate through their union officials. A National Industrial Conference, called by President Wilson in 1919 to work out a national system for the adjustment of trade disputes, broke up because of the employers' refusal to accept the principle of collective bargaining through union officials. The employers' group insisted on negotiation with the workers in each individual plant through representatives chosen "by and from them." The joint council plans of Great Britain, Germany, and other countries are debarred from the United States until trade unionism has established itself as a normal feature of the industrial system.

Employee representation. Many large employers in the United States, who refuse to negotiate with the trade unions, have set up conciliation machinery within their own industries. Though differing in detail, these plans embody the common principle of employee representation on those councils of the enterprise which control matters of policy most closely affecting the welfare of the wage-earners. They also agree in their essential opposition to the principle of trade unionism. Trade unionism contemplates a horizontal division of society into two opposing classes which cut across the boundaries of the separate business establishments. It views the interests of these two classes as inevitably antagonistic to each other. They are organized for warfare but adopt negotiatory devices, of much the same nature as the diplomatic systems of nations, to handle the joint affairs of the union and employers' association during periods of truce. In the philosophy of trade unionism industrial peace is to be secured by an equipoise of power, which arouses in each party a wholesome respect for the organized force of the other. Conciliation machinery, such as has been described in the preceding pages, is considered of service only because it is backed up by the weapons of the negotiators. Employee representation, on the other hand, intends an organization within each industrial unit separately. It presupposes a mutuality of interest between the employer and the worker and is devised to increase the competitive power of the enterprise in its struggle with other business units by promoting harmony and

efficiency. The shop committee or council, made up of equal representatives of the wage-earners in the plant and the management, supplies a point of contact between these two interests where grievances may be adjusted and terms of employment fixed by agreement.

American experience with employee representation. Employee representation is a distinctively American development in the field of industrial relations. It grew up, during and after the World War, as an offset to declining labor efficiency and a safeguard against the spread of trade unionism with its program of the closed shop. Some of the largest of our business units have applied the system, in forms adapted to the peculiarities of their enterprises, with apparent success. Following a violent struggle in the mines in 1915, the Colorado Fuel and Iron Company set up a complete system of employee representation with a structure of joint committees reaching from the individual mine to a general council at the top, which handles matters of concern to the enterprise as a whole. The Standard Oil Companies and their subsidiaries, the Goodyear Rubber Company, the Pennsylvania Railroad Company, the General Electric Company, the Bethlehem Steel Corporation, and the great meat packing companies are among the large concerns which have adopted similar plans. Early in 1925 it was estimated that approximately eight hundred different business units in this country, employing upwards of a million and a half workmen, had installed some system of employee representation.

Employee representation is intended only incidentally to serve as a conciliation device. Its primary purpose is to promote efficiency in production; in other words, it is essentially a method of *labor management*. On the basis of their brief experience with the system, American industries are unanimous in their verdict as to its success. It has lowered costs and increased output. Incidentally it has promoted industrial peace, by providing a mechanism for the clearance of small grievances as they arise in the routine operation of the shop and for winning the prior consent of labor to changes in working conditions which might otherwise have

precipitated a fight. When it was found necessary to reduce money wages after the panic of 1920, concerns with employee representation systems were able to make the change without much trouble; whereas in the closed shop industries the move precipitated a series of strikes. It is still too early to say however whether this difference was not more due to the greater bargaining weakness of the workers who had no support from national trade-unions than to their conviction as to the justice of their employers' demands.

The system of employee representation contains weaknesses, which are much stressed by the leaders of organized labor: (1) representatives of the workmen lack courage to defend their interests because their jobs depend on the goodwill of the employer, (2) these representatives, being ordinary wage earners continuously employed at their jobs, can acquire no knowledge of the conditions of the trade and no expertness in negotiation, (3) final control rests in all cases with the management, which may veto the action of the shop committee and abolish the whole system at will, (4) the system binds the wage-earners to their individual enterprise, subjecting them to its hazards and depriving them of the collective support of other workers in case they are driven in the last resort to strike in self-defense.

Mediation. Mediation may be viewed as the first step in the direction of acknowledging the right of the public to a voice in the affairs of employer and employee. It involves a friendly and informal effort by some individual or agency not party to the dispute to bring the disputants together for the purpose of working out a peaceful settlement. The function of the mediator is diplomatic rather than judicial. That informal, non-coercive intervention of this sort is so frequently successful may be attributed to human nature. "When once a difference has been accentuated and, still more, when it has developed into an open conflict, both sides are apt to be striving for 'mastery' as well as for the particular object of the dispute. They stand to lose dignity as well as money; and, consequently, their obstinacy will be much greater than the material point alone can justify. Not only is this, as a

matter of fact, so, but it is frequently known to be so by the parties themselves. They will often have considered some matter worth the *chance* of a rupture but not worth the certainty of one. Hence when the rupture actually arrives, all that is needed is some device for facilitating withdrawal without undue loss of dignity from a position assumed for purposes of bluff. Even if, in the earlier stages of a conflict, no way seems acceptable, a point is sure to be reached sooner or later when one party will be willing to yield if it can 'save its face' in doing so. Hence the opportunity for the 'good offices' of the mediator." [1]

Mediators may be self-appointed men of prominence in the social or political world who come forward at the critical moment to volunteer their services. In England the prime minister of the day or some eminent man of noble rank has succeeded at times in mediating disputes when all attempts at direct negotiation had failed. In America mayors of the cities, governors of states, and even the President have performed this function with success. The eminence of such a mediator "flatters the disputants with a sense of their own importance" and inclines them to make concessions as a personal favor to him which they would refuse to make to an opponent.

(A) **Official mediation.** Many countries however have provided regular machinery for mediation in the form of standing boards of public officials. In England, under the Industrial Courts Act of 1919, the Ministry of Labor is endowed with mediatorial power, which may be used on occasion with the consent of both parties.

In the United States the Erdman Act of 1898 provided that the Chairman of the Interstate Commerce Commission and the Commissioner of Labor should offer their services as mediators in disputes on the railways. Power to act as arbitrators was also given to these officials. Between 1898 and 1912 forty-eight applications for mediation and arbitration were made to this board, the majority of the cases being settled by mediation. When the Newlands Act superseded the Erdman Act in 1913, provision was made for a Commission of Mediation and Arbitration to be

[1] Pigou, *Economics of Welfare*, Macmillan, 1920, p. 386.

appointed by the President for a term of seven years. Of the seventy-one cases in which the services of the Commission were employed between 1913 and 1917, most were settled by mediation. The failure of the Commission to settle a threatened national strike in 1917 marked the end of the usefulness of this agency. The law of 1913 which created the Department of Labor empowered the Secretary of Labor to act as mediator. Between 1913 and 1921, 3,500 cases were adjusted through the good offices of the Department. Agencies created during the World War, The President's Mediation Commission of 1917 and the War Labor Board (1918–1919) under the chairmanship of Ex-President Taft, were strikingly successful in preserving peace during this trying period. Most of the states have added mediation to the list of duties of their industrial commissions or departments of labor.

In France and Belgium the ordinary local officials of the government are empowered to act as mediators. The experience of the industrial world as a whole with this informal procedure is sufficient to prove its usefulness as an agency of peace. Its danger, as pointed out by Professor Pigou, lies in the fact that the availability of mediatory officials may prevent the development of conciliation machinery within the industries themselves, which is a more stable and effective negotiation device than mediation.

Voluntary arbitration. Next in order in the direction of public settlement of the industrial conflict is the voluntary submission of disputes to an impartial umpire or arbitrator. Voluntary arbitration usually presupposes an existing trade agreement between organized labor and the employer which calls for the final determination of disputed points by a third party. Consent of both parties must be obtained before a cause can be submitted to an arbitrator, and the award is binding on neither party, since it is non-enforcible at law. However, when a trade agreement contains an arbitration clause, it is extremely difficult for either party to refuse later to submit a cause in dispute without loss of public support, and continued practice of arbitration gradually builds up a moral sanction which has great force in controlling conduct. The acceptance of the award is voluntary, but the very

facts that all previous steps have also been voluntary and that the arbitrator has been chosen by both sides because of his impartial character give the award a sanction which is often more binding than a judgment based on force. Experience has shown that few unions or employers repudiate arbitration awards, provided the arbitrator has not overstepped the bounds of his mandate.

Arbitration agreements and procedure. The arbitrator, like the mediator, may be either a private individual or board or a public officer or public board. The typical unofficial machinery in this country consists of a board of three, two representing employers and workers respectively while the third is an impartial chairman selected by these representatives or named in the trade agreement. In such cases the decision almost invariably rests with the chairman. Yet the other members are not without their uses; they serve as channels of information to the arbitrator, and their presence on the board adds to the influence of the decision when it is announced. Many large industries give full-time employment to an impartial umpire with the duty of interpreting disputed points in the trade agreement and adjudicating small causes of controversy as they arise. The development of trade agreements between national unions and large employers or employers' associations has created a new profession of arbitrators, men skilled in the technicalities of the trade, whose position is akin to that of the judge interpreting law. The Hart Schaffner & Marx agreement covering a large section of the Chicago clothing market provides for the services of such an official. The Amalgamated Clothing Workers, the United Garment Workers, the Journeymen Tailors Union, the Ladies' Garment Workers, and the United Cap and Hat Makers have so generally incorporated arbitration provisions in their trade agreements that approximately 400,000 workers in the clothing trades are now embraced in this machinery for preserving industrial peace. In the printing trades, the International Typographical Union and the publishers recognizing the closed shop have a ten-year national agreement providing for arbitration of all disputes. In the building trades these arbitration devices are local in nature, covering each city

market independently; an example is the Joint Arbitration Board of Chicago. All of these agreements stipulate that there shall be no strike or lockout pending arbitration proceedings. It will be noted that this use of arbitration by prior agreement is feasible only in highly organized industries where trade-unionism and collective bargaining are established principles. All unions which are no longer fighting for recognition favor arbitration. The system is thoroughly established throughout the industrial world wherever trade-unionism flourishes.

Official arbitration. Almost all governments have set aside public officials or boards to act as arbitrators on request of the parties to a dispute, either as an incidental feature of their duties, or as their sole function. Officials or boards empowered to offer their services as mediators may also in most cases serve as arbitrators. Virtually every one of the American states has made some such provision for voluntary arbitration. The Erdman and Newland Acts, referred to above, provided that the designated mediation officials were to act as arbitrators of disputes on the railways upon consent of both parties if conciliation failed to compose the controversy. On the whole these *ex officio* arbitrators have had little success either in this country or elsewhere. Both sides are inclined to distrust the impartiality and the capacity of men who owe their positions to political fortune. An improvement on the system is the appointment of a panel from which arbitrators may be chosen by the two parties, rather than the designation of a specific official; this is done by the Industrial Courts Act of 1919 in England. A somewhat similar variation is found in the laws of some twenty American states, which provide for the appointment by the Governor of representatives on the Board of Arbitration from nominees of employers and workers.

Results of voluntary arbitration. As long as the arbitrators do not much concern themselves with principles of abstract justice and right, voluntary arbitration has proved successful in promoting peace in industry. In simple cases, such as the interpretation of doubtful clauses in an existing trade agreement, the arbitrator has little difficulty in reaching an acceptable decision by applying

fairmindedness and common sense to the problem. But when the controversy involves some change in the *status quo* — such as an increase or decrease in wages, alteration in hours, etc. — the most effective procedure is for the arbitrator to adjust his award to the relative strength of the two parties. A shrewd appraisal of the power of the two organizations and of their determination to fight in defense of their positions gives the arbitrator a basis on which to forecast the probable outcome of a strike or lockout. He then attempts to phrase a decision which will anticipate this result. Wages, hours, and other industrial conditions fixed by arbitration are not deductions from economic law or from premises of public interest, but compromises between two contending forces. The document which promulgates the award is often embellished with reference to general economic principles, human welfare, costs of living, etc.; but these are usually intended to make the decision palatable to the defeated party and to win public support. The justification of the arbitrator's activities is not that it ensures the operation of justice and right, but that it saves employer, worker, and the public from the losses involved in industrial warfare, while yet according to the potential victor the spoils of war. Experienced private arbitrators will freely admit that they are guided by these "practical" considerations. Public boards and officials are prone to seek a higher plane on which to base their decisions, with the result that arbitration machinery frequently breaks down under their administration. It is not to be expected that an organized group will continue to favor this method of settlement if its net result is to deprive it of the advantages of a strong bargaining position.

Introduction of compulsion. There are several transition stages between such voluntary arrangements as we have described and a fully developed system of coercive public intervention. One intermediate stage is represented by laws which make submission of disputes to arbitration compulsory, while leaving the acceptance of the award voluntary. An illustration is the Transportation Act of 1920, creating the Railroad Labor Board, which has been discussed in another place. Another stage is reached when at-

tempts are made to give the award binding force, after arbitration has occurred through mutual consent of the two parties. In an informal way this is sometimes done by requiring the parties to an arbitration agreement to deposit sums of money, which are forfeited in event of refusal to accept the decision of the arbitrator. The laws of about a dozen of our states provide for punishment, by fine or imprisonment, of a recalcitrant party who has previously agreed to arbitration by public officials. These laws have proved impossible to enforce, and their presence on the statute books has acted as a deterrent to voluntary arbitration. It is impossible for us to attempt even a summary of these various expedients which lie midway between voluntary and compulsory arbitration. But something must be said regarding one form of partial public coercion which has attracted much attention in the field of industrial relations — the system of compulsory investigation.

Compulsory investigation. The first announcement of the right of the public to knowledge of the facts surrounding industrial disputes and the first provision of machinery for the collection and publication of this information were contained in certain state laws in the United States of America. It is the experience of Canada, however, which has brought the device of compulsory investigation into prominence. A severe coal strike in 1906 led to the passage of the Canadian Industrial Disputes Act in the following year. The Act applies only to public utilities, transportation agencies, and mines. It provides that any change in wages and hours must be preceded by thirty days' notice. If the change is unacceptable to one of the parties, no strike or lockout may occur until after investigation of the facts by a board appointed by the Minster of Labor and public announcement of a decision by the board. Fines may be imposed for illegal strikes or lockouts during the investigation period. After the announcement of the award, however, both sides are at liberty to resort to a trial of strength. The Act has not stopped strikes. During the first ten years of its life, down to the end of 1916, 222 strikes and lockouts occurred, involving over 100,000 workmen and a total loss of 4,838,647 working days; 204 of these disputes were illegal under

the Act as occurring before or during investigation; 18 were repudiations of the award of the investigation commission.[1] In 1918 there were 196 strikes and lockouts. In the following year 296 strikes involved a loss of 4,000,000 working days.[2] The penalty features of the law have been practically inoperative. During the first ten year period only eleven persons were prosecuted, although upwards of 80,000 had broken the law, and only $1,660 were collected in fines as compared with $150,000,000 as a maximum legally due. On the positive side the system must be credited with the settlement of a great many disputes which might otherwise have resulted in open breaks. Its success seems to depend chiefly on the attitude of organized labor. Because the railroad unions favor the system, it has succeeded exceptionally well in this industry. The same is true of the eastern section of the mining industry.

A system very similar to the Canadian was set up in the State of Colorado by the Industrial Commissions Act of 1915. This Act also was passed at the end of a disastrous strike in the mines. The law applies to all industries except in domestic service, in agriculture, and those employing less than four men. It copies from the Canadian law the requirement of thirty days' notice of change in working conditions, the prohibition of strikes pending investigation, and the public announcement of a non-mandatory award by the Industrial Commission. The period since 1915 has been one of industrial peace in Colorado; but it is impossible to say to what extent this has been due to the operation of compulsory investigation and to what extent to a system of conciliation set up by the mine operators at the same time.

Appraisal of compulsory investigation. The strength of a system of compulsory investigation lies in the facility with which it arouses public opinion in support of that party to the dispute who agrees to accept the findings of the Commission. Fear of public condemnation promotes a willingness to compromise which results in the settlement of many potential causes of strike by

[1] *Monthly Labor Review*, Sept. 1917, p. 11.
[2] *Ibid.*, July, 1921, p. 24.

conciliation procedure, often without outside aid. Its weakness lies in the fact that it is almost universally condemned by organized labor. This is due principally to the fact that the enforced postponement of hostilities, which is the essential feature of the system, works to the detriment of labor and in favor of the employer. The success of a strike depends in large measure on the complete tie-up of the industry. A waiting period gives time for the assembling of strike-breakers, the shifting of orders to other concerns, and similar expedients, which reduce the employer's loss. The strikers gain nothing by being compelled to wait and are in danger of losing the morale and *esprit de corps* which are vital to success in a protracted struggle. For this reason the system is viewed with suspicion by organized labor as a device to strengthen the hand of the employer. Irritating delays in the procedure of official commissions and a distrust of the fairness of mind shown by the ordinary political appointee are other causes of the opposition of labor unions.

The case for compulsory arbitration. Interference by representatives of the public in the relations of worker and employer reaches its extreme form when strikes and lockouts are entirely forbidden and provision is made for compulsory arbitration with an award enforced by governmental power. Many are convinced that compulsory arbitration of all trade disputes must be the inevitable outcome of the progress of law. Historically the growing interest of the state in the preservation of public order has led to a widening of the scope of law to cover actions of individuals and social groups which were previously considered purely private in nature. Revenge, the blood feud, and retaliation have been suppressed, as the state has gathered into its hands the duty of administering justice and protecting the private rights of its citizens. In advanced nations the settlement of disputes by private contest, in all phases of human relationship except the relation of employer and worker, has been discarded in favor of the judicial process sanctioned by the force of government. That the extension of law to this important field has been delayed is attributed, by the proponents of compulsory arbitra-

tion, to the fact that public interest in industrial relations has been of late development, arising only after the establishment of the factory system and the growth of trade-unionism. Recognition of the right of the public to a voice in the settlement of trade disputes, it is argued, must lead in time to a disarming of capital and labor for private warfare. This argument from analogy is impressive; but, as we shall see, it carries with it implications of grave social importance which are not always fully understood by those who favor compulsory arbitration.

Experience in New Zealand and other states of Australasia. New Zealand was the first country to adopt a system of compulsory arbitration. The first law was passed in 1894, after a period of bitter strikes. Since that date there has been an unbroken history of compulsion in New Zealand. The existing system has as its basis a series of regional conciliation commissions covering the entire island, each commission under the direction of a salaried expert. Attempt is made to settle all grievances by voluntary adjustment with the aid of these commissions. Failing this, an appeal from one of the parties takes the case before a Court of Arbitration, which has final jurisdiction over all disputes. Until 1914 strikes and lockouts were not illegal unless application had been made for an award by the Court, but since that date all strikes have been prohibited in the case of unions registered under the Act. Registration by the unions is voluntary, and those which have not registered are not denied the right to strike. However the ease with which an entire labor group may be registered by a minority of its members gives the system of compulsion wide sway. A law of 1913 set up a scheme of compulsory investigation, similar to that of Canada, applicable to unregistered workers. This new law prohibits strikes during the investigation period.

The Court of Arbitration consists of a president, taken from among the justices of the Supreme Court, and two other members chosen from nominees of the registered unions and the registered employers' associations respectively. The Court has judicial power to compel the attendance of witnesses and the submission of books and documents. Violation of awards is punishable by

fine, and each award holds good for three years unless superseded by a new judgment of the Court, even though the union to which it applies has cancelled its registration in the meantime. Awards issued in cases affecting a single enterprise may be made applicable to the entire industry, thus equalizing competitive conditions.

Australia has developed a system of compulsory arbitration patterned after that of New Zealand. The Federal Commonwealth adopted a law in 1904, which prohibited strikes in interstate industries and set up a Court of Arbitration. Each of the different states has adopted some system of compulsion. Since 1901–1902 New South Wales and Western Australia have had compulsory arbitration laws, which apply to all industries and all workers without exception. Queensland and South Australia adopted a similar system in 1912. Victoria has a system of wage boards for the different industries, whose principal function is to guarantee "living" wages to the workers but which operates also for the adjustment of agrievances.

Compulsory arbitration in Kansas. Outside of Australasia one attempt has been made to apply the principle of compulsion. The State of Kansas in 1920 passed a law prohibiting strikes and lockouts in specified industries and creating a Court of Industrial Relations to settle all causes in dispute. The industries affected were those engaged in the manufacture of clothing and food products, mining, transportation, and public utilities. The Court was authorized to intervene in disputes arising within these industries, either on its own motion, on appeal of one of the parties, on petition by a small group of taxpayers, or at the order of the attorney general of the state. Its powers extended to the fixing of wages, hours, and working conditions, and, to prevent suspension of operations, it was empowered to take over any enterprise, guaranteeing a fair return to capital and labor. Its awards were mandatory, sanctioned by penalties of fine or imprisonment. Early in 1925 however the Court was shorn of most of its power by a decision of the Supreme Court of the United States, which declared unconstitutional the compulsory arbitration features of the law as applied to working time and by implication overthrew the whole

process of compulsory arbitration whenever its effect was "to compel owner and employees to continue the business on terms not of their own making." Just prior to this decision, the Court had been abolished by the legislature, its powers being transferred to a public utilities commission. But the position taken by the Supreme Court has apparently prohibited all further attempt to apply the principle of compulsion to privately owned enterprises in this country, except in cases of emergency in industries clearly affected with public interest, as long as our Constitution contains a clause guaranteeing "liberty of contract."

Judgment of the Australasian experience. Any inferences regarding the practical operation of compulsory arbitration must be drawn from the experience of Australasia, where this method of promoting industrial peace has had thorough trial. For a time the world was inclined to accept the evidence of this experience as proof of success. New Zealand was known until 1905 as a "land without strikes." But the subsequent history of the system has shown that strikes cannot be prevented by compulsion. Since 1906 New Zealand has experienced frequent and bitterly contested struggles in the field of industrial relations. Australia has had a similar experience. The attempt of the Commonwealth to stiffen the penalty features of the law by imprisoning strikers precipitated such widespread revolt that the severity of the penalties had to be modified.

It is becoming apparent that the early success of these laws was due chiefly to the contentment of the workers with a system which operated continuously in their favor. The Courts of Arbitration favored the principle of unionism, even to the extent of compelling employers to accept the closed shop. Early awards resulted in continuous advance of wages and shortening of hours, thus conceding to labor larger material gains than could have been obtained from the process of bargaining alone. There were no strikes, because the causes of strikes were absent. The underlying reason for this result favorable to labor was the rapid and sustained material progress of Australasia during this early period of its industrial development, which made possible a continued

increase of the wage income without encroaching on the necessary rewards of the other factors of production. The rate of progress was retarded during the first decade of the twentieth century, and when it became impossible to grant the demands of the workers rebellion against the system arose immediately. Today there is a formidable " radical " wing of the organized labor movement, which repudiates compulsory arbitration and strikes against the system as a matter of principle.

Inherent defect of compulsory arbitration. The most serious defect of compulsory arbitration however is not its record of incomplete success but its essential incompatibility with the institutions of a capitalistic society. When a public body undertakes to fix wages, hours, and other terms of employment, it is ultimately compelled to discard economic principles in favor of a determination based on moral or ethical considerations. Free play of economic forces apportions an income to labor, as to the other productive factors, which measures its marginal indispensability in society's industrial structure. Boards of arbitration are forced by popular pressure to fix wages with the intent of guaranteeing a decent or wholesome standard of living to the workers regardless of their competitive weakness. No single concession to the workers will keep them content for any length of time. Like all other social groups, they press for a continued improvement in their economic status. The practical difficulty of enforcing the award puts pressure on arbitration officials to placate the labor side of each controversy. Penalties can be applied with ease to the ownership faction, whereas fines levied upon scattered thousands of workers cannot be collected without great trouble and expense. Consequently the standard of living used as a basis for wage settlements must be raised to keep pace with the expanding demands of labor.

But the determination of one type of income according to arbitrary standards of justice and right necessarily involves a similarly authoritative determination of other types of income. When wages are raised out of compassion with the workers or from motives of political expediency, a limitation is set, by im-

plication, on the shares of the social income distributed as interest and profits. A board of arbitration, setting out with the simple object of finding a practical adjustment of isolated disputes in individual enterprises, is confronted in time with the task of determining what is a "fair" rate of return for the infinite variety of services performed by land, capital, and management, as well as by wage labor. There is no final solution of this problem in a society where reliance is placed on private initiative to supply requisite amounts of wealth and of labor and to undertake the direction of the productive process.

The opening wedge of socialism. Considered from the standpoint of its essential economic implications, compulsory arbitration is correctly viewed as the "opening wedge of socialism."[1] It involves a gradual interference with property rights, freedom of contract, and private management, whose ultimate consequence is not always fully understood by those who favor it. The next step after fixing a rate of wages which is not acceptable to the workers is to compel free men to perform labor at that rate. If interest returns are lowered to a degree which retards investment, the government is confronted with the necessity of maintaining and improving capital equipment. The strike weapon cannot be taken from the workers unless the employers are deprived of the correlative weapon of discharge. This and other forms of interference with the rights of management, coupled with a possible restriction of profits, may proceed to the point of driving private enterprisers out of industry. Such a contingency will place on the government the obligation of ownership and management of enterprises whose collapse would seriously affect public welfare. It must be admitted that these implications of a system of compulsory arbitration are as yet only theoretical; they have not materialized in the experience of the countries employing that system. But compulsory arbitration has been brought to a decisive test only within recent years, when the slower pace of material progress has forced into clear relief the conflict between what is desirable, from the humanitarian viewpoint, as a reward

[1] Taussig, *Principles of Economics*, Vol. ii, p. 349.

for labor and what is economically possible without injury to other interests in production. There is now a clear tendency in these countries which employ compulsory arbitration for the distribution of income to become an issue between opposing political parties which struggle for mastery of the machinery of arbitration.

The future of the industrial conflict. Though it is unsafe to venture a prediction regarding the future of a rapidly changing industrial order, certain conclusions may be drawn from the world's experience with the various methods of solving industrial disputes which we have discussed. It is clear that organized labor cannot be forced to surrender its right to strike as long as the powers of the enterpriser remain uncontrolled. Public opinion will not support a system of industry in which men without power to resist are employed for private profit and under the command of the profit-maker. In England and America members of the police and fire departments of the cities have been denied the right of striking; but this is practicable because they are public employees who accept these forms of employment in preference to an abundance of opportunities in other lines. The suggestion that an anti-strike clause be introduced into the Transportation Act of 1920 aroused such menacing opposition from labor that the bill had to be modified. It is a safe prediction that strikes cannot be prohibited even in such essential industries as transportation and mining unless the government is prepared to take over ownership and operation of these industries. This would mean that competitive control of prices, wages, and profits had been discarded in favor of some non-economic form of control.

No type of public intervention has as yet proved as effective in promoting industrial peace as have voluntary arrangements set up within industry by mutual consent of employer and worker. Conciliation, mediation, and voluntary arbitration are accomplishing much in the way of smoothing out trade disputes. Machinery for public intervention in the last resort is most useful when it is carefully designed to promote these informal devices for negotiation and settlement. Attempts to apply compulsion instantly

arouse the suspicion and antagonism of labor. If they are persisted in, despite this antagonism, their ultimate effect is to force organized labor into politics, with a program of far-reaching social change which may mean the end of the present system of industry. But voluntary arrangements, though effective to a degree, have done little more than mitigate the menace of the strike. Hope for the future appears to lie along the line of a readjustment of the relationships of employer and wage-earner tending to remove the natural clash of interest between these two factions, rather than in expedients for settling differences after they have arisen.

Suggestions for Further Reading on Part VII. Labor Problems

I. General

Adams, T. S. and Sumner, H. *Labor Problems.* New York, 1905
Watkins, G. S. *Introduction to the Study of Labor Problems.* New York, 1922
King, W. L. K. *Industry and Humanity.* Boston, 1918
Rowntree, B. S. *The Human Needs of Labor.* London, 1921

II. Unemployment

Beveridge, W. H. *Unemployment.* London, 1909
Feldman, H. *Regularization of Employment.* New York, 1925
Klein, P. *The Burden of Unemployment.* New York, 1923
National Bureau of Economic Research. *Business Cycles and Unemployment.* New York, 1923

III. Accident, Sickness, Old Age, and Labor Legislation

Lauck, W. J. and Sydenstricker, E. *Conditions of Labor in American Industries.* New York, 1917
Goldmark, J. *Fatigue and Efficiency.* New York, 1912
Eastman, C. *Work Accidents and the Law.* New York, 1916
Epstein, A. *Facing Old Age.* New York, 1922
Commons, J. R. and Andrews, J. B. *Principles of Labor Legislation.* New York, 1920
Pigou, A. C. *Economics of Welfare.* Chapters 13–17. New York, 1920
Downey, E. H. *Workmen's Compensation.* New York, 1924
Seager, H. R. *Social Insurance.* New York, 1910
Woodbury, R. M. *Social Insurance, an Economic Analysis.* New York, 1917
Cohen, J. L. *Insurance by Industry.* London, 1923

IV. Trade Unionism

Webb, B. and S. *Industrial Democracy.* Part iii. New York, 1920
Hoxie, R. F. *Trade Unionism in the United States.* New York, 1917
Cole, G. D. H. *The World of Labor.* London, 1918
Perlman, S. *History of Trade Unionism in the United States.* New York, 1922
Marot, H. *American Labor Unions.* New York, 1914
Commons, J. R. (Editor) *Trade Unionism and Allied Problems.* New York, 1921
Fitch, J. A. *The Causes of Industrial Unrest.* New York, 1924
Wolman, L. *Growth of American Trade Unions, 1890–1923.* New York, 1925

V. Conciliation, Mediation, Arbitration

Gilchrist, R. N. *Conciliation and Arbitration.* Calcutta, 1922
Barnett, G. E. and McCabe, D. A. *Mediation, Investigation, and Arbitration.* New York, 1916
Clark, V. S. *The Labor Movement in Australasia.* New York, 1906
Higgins, H. B. *A New Province for Law and Order.* London, 1922
National Industrial Conference Board:
Arbitration and Wage Fixing in Australia. Boston, 1918
Canadian Industrial Disputes Act. Boston, 1918
Conciliation and Arbitration in New Zealand. Boston, 1919

PART VIII

REFORM OF THE ECONOMIC SYSTEM

CHAPTER LV

PROFIT SHARING, CO-PARTNERSHIP, AND COÖPERATION

There are many who question the soundness of the wage system as a form of industrial organization. Its failure to stimulate the efficiency of labor, the growing friction between employers and workers, and the failure of all devices to solve the problem of the industrial conflict have called into existence various types of reform which aim at a more or less complete reorganization of human relationship within the productive unit. The most extreme of these reform movements are represented by the socialist programs to be studied in the following chapter. Other programs are less revolutionary, in that they intend to leave unchanged our basic institutions of private property, competition, and individual enterprise. Profit sharing, co-partnership, and (on the whole) coöperation are of this more conservative nature.

Profit sharing. The profit sharing movement had its origin almost a century ago. Its striking success in the Léclaire establishment in Paris, where it was introduced in 1838, and in the Godin Familistére, which adopted it two years later, attracted much attention among progressive employers in Europe and Great Britain. During the middle decades of the nineteenth century many experiments along this line were begun. For a time profit sharing was hailed by those with a bent toward reform as a complete and final solution of the labor problem. This early enthusiasm was short-lived. With a few notable exceptions, the first profit sharing schemes collapsed, leaving those who had pinned their faith to them disillusioned and others wary of the risk of failure. But now, within our own generation, interest in profit sharing has revived. Many large employers have introduced the system with complete success, and it has become again a subject of much discussion and

study among those concerned with solving the problem of industrial relations. This renewed interest in a system which after much experiment had been condemned as a failure is due in large part to the spread of trade-unionism with its threat of encroachment on the powers of management in industry. Modern systems of profit sharing are successful chiefly because the business world has learned the limitations of this method of organizing the relation of employer and worker and has avoided failure by not attempting too much.

Prerequisites of true profit sharing. The essential principle of profit sharing is simple. It introduces a new basis of distributing the corporate income. After normal returns to the ownership and labor factions have been paid, the residual profits — ordinarily accruing entirely to the owners — are divided between the two according to some predetermined scale of sharing. Many arrangements which are not true profit sharing parade under that name. Some employers for example have the habit of bestowing gratuities on their employees at Christmas time or bonuses at the end of a prosperous year. The purpose of these pseudo profit sharing schemes may be the same as that of true profit sharing; though usually it is not. "Such methods are of questionable value viewed as instruments either for the promotion of efficiency or of social welfare. The arbitrary decision of the management, made after and not before the extra exertion, robs the plans of a large part of their stimulating power. It is not to be expected that men will concern themselves to increase profits and keep down costs, when they have no knowledge as to what portion of the profits, if any, they will receive. From the standpoint of social betterment, such arbitrary distributions are also to be condemned. A widespread use of profit sharing of this character would inevitably work undesirable results among members of the employed class, by making them more dependent on the bounty and charity of employers than on their own initiative and bargaining power."[1]

True profit sharing must be free from any taint of charity or philanthropy, a business arrangement between voluntarily con-

[1] A. W. Burritt, *et al.* *Profit Sharing, Its Principles and Practice*, 1918, p. 11.

tracting parties. The plan must be definite and clearly understood in advance and must bestow on labor a right to automatic increases in income under certain conditions. These contingent earnings of labor must be in excess of the market rate of wages. Profit sharing will not be effective unless the workers embraced in the scheme are guaranteed an income equivalent to that of other workers in their group and given, in addition, the chance of sharing in the prosperity of their enterprise.

Typical plan of profit sharing. The precise basis of sharing residual profit between workers and owners varies with different profit sharing plans. Some such arrangement as the following is typical. After payment of costs including wages, five per cent interest is paid on the capital stock and the remainder of the net profit is then divided between stockholders and workers in the ratio that total face value of capital bears to total wages paid during the year. In some concerns the residual income is divided in the ratio of total interest (rather than capital value) to total wages; this obviously increases the amount of profits accruing to labor. It is customary to subdivide labor's share of profits among the individual workmen in proportion to their normal wages. This is a recognition of the fact that market rates of wages are an index to the relative productive importance of different groups of workers.

Profit sharing may lead to producers' coöperation. With very few exceptions, profit sharing plans provide that the workers' share of the residual income must be reinvested in the enterprise, being distributed in stock or other certificates of ownership. This is a natural arrangement in view of the purposes of the profit sharing system. A gradually accumulating share of ownership in the hands of the workers will tend more completely to merge their interest with that of the employer than would a mere increase in their annual money income. It gives them in time a double stake in the prosperity of the enterprise: they share in the guaranteed interest income as well as in the current residual profits.

This feature of the system has the ultimate effect of giving profit sharing a social significance quite different from that with which it

sets out. Successful experiments in profit sharing usually pay the workers a minimum share in the profits equal to ten per cent of their wages, and sometimes this share rises as high as twenty-five or thirty per cent. These percentages of the total pay roll of a large enterprise represent a considerable amount of money which, being reinvested year by year, may in time give labor a surprisingly large fraction of ownership. This tends to grow as the years pass, unless measures are taken to maintain the proportion of ownership in the hands of other types of investors. The absolute amount of money accruing to the workers almost invariably exceeds that accruing to the owners, because the pay roll of the typical business enterprise is larger than its capital investment. If now both groups of sharers are obliged to invest their part of the profits in the company's stock, labor's investment will grow faster than that of other investors. This process may result at length in transferring control over the concern to the workers, thus causing the profit sharing scheme to culminate in a system of producers' coöperation — a system which raises social problems of quite a different order. This result is not inevitable, since it requires that the original group of workers and their descendants remain permanently attached to the enterprise. But since one of the effects of profit sharing is to reduce the turnover of labor, this event is by no means improbable. We shall examine the relationship of profit sharing to co-partnership and to producers' coöperation in later sections of this chapter.

Examples of profit sharing. The N. O. Nelson Company of St. Louis began a tentative system of profit sharing in 1886, which will serve as a typical modern example of profit sharing. With some changes in detail, this system has been successfully continued down to the present time. After payment of the normal rate of wages, and six per cent on the capital, the remaining net income is divided between owners and workers in the proportion that total wages bear to total capital. The workers' share is invested in the stock of the company. This stock is withheld from its owner for a three-year period, its dividends being credited in the meantime to the purchase of more stock. An employee leaving the concern

before the expiration of the three-year period is given cash equivalent to his past profits but loses the dividends which have been accumulating on his stock. The workers are prohibited from buying or selling stock issued to them, though the company stands ready to repurchase the stock of employees in need of financial assistance. Since the inauguration of the plan, the capital stock of the concern has increased from $100,000 to $2,000,000, and the number of employees from two hundred to a thousand. At present one third of the capital is owned by the workers, and this fraction is steadily mounting.

In 1916 the Bureau of Labor Statistics reported sixty American concerns as using some form of profit sharing. The practice of adopting a profit sharing system in conjunction with employee representation has since increased this number. Among the more important business establishments in the profit sharing group in this country are the Proctor and Gamble Company, whose employees own nearly $2,000,000 worth of the Company's stock, the Dennison Manufacturing Company whose profit sharing employees control the voting power in the concern, Sears Roebuck and Company, the Henry Ford Company, the Boston Consolidated Gas Company, the Farr Alpaca Company, the William Filene's Sons Company. In other countries, the most outstanding examples are the Maison Léclaire and the Godin Familistére of France. The former has had successful experience with profit sharing since 1838, and the concern is now entirely owned and controlled by the workers. The latter adopted its plan of profit sharing in 1840, and this plan also has developed into a system of producers' coöperation. In England Lever Brothers, the great soap manufacturers, the Taylor Company, producers of woollen cloth, Armstrong, Whitworth and Company, makers of battleships, and many of the largest gas and light companies are notable examples of successful profit sharing.

Criticism of profit sharing. Any appraisal of the merits of profit sharing must grant at the outset that no far reaching social changes are to be expected from this device. It will be adopted only on the initiative of the employer and because it holds forth promise

of material advantage to himself. The employer retains control over the plan and can discard it at will. Only when a profit sharing scheme has ripened into a system of producers' coöperation does it cause any fundamental change in the wage system. This eventuality, though always a potential feature of profit sharing, will never materialize unless the scheme is under the direction of an employer convinced of the undesirability of the wage system. No program of social reform which relies for its success upon the altruism of individual business leaders can hope to accomplish much in a world dominated by self-interest. Profit sharing must be accepted as a plan with definite and narrowly practical objectives. Its purpose in the modern world is twofold: to stimulate the productive efficiency of the workers and to promote harmony of industrial relations. The latter is subordinate to the former of these purposes; the employer is interested in industrial harmony as a means toward greater efficiency.

Limited usefulness in promoting efficiency. As a device of management to promote efficiency, the scope of profit sharing is limited. In the nature of the case it can be effective only when the character of the occupation is such that a clear relationship exists between the workers' faithfulness and industry and the profits of the concern. The individual employee must foresee definite returns from increased industry on his part. But this relationship is by no means clear in the case of ordinary types of labor in a large business unit. The profit of the concern is strongly affected by factors quite beyond the workers' control. The financial policy of the enterprise, the behavior of markets, changes in international economic relations may cause net loss to the business despite the best efforts of the wage-earners. Even aside from these extraneous forces, the effect upon profits of greater industry on the part of the individual laborer may be impossible to detect. Refined division of labor makes the connection between the productive efficiency of a single wage-earner and the profits of the enterprise remote and indirect.

The limitations of profit sharing as a device for promoting efficiency have been summed up in two formulas by a careful student

of the modern experience. First, *the efficiency of profit sharing varies directly with the rank of the worker.* The salaried and higher wage groups respond most noticeably to the appeal of the system. The reason is not far to seek; in these higher groups the effort of each individual counts for much in the income of the enterprise. Recognizing this fact, the Dennison Company limits its profit sharing plan to those whose annual wage is $2,500 or more. The Studebaker Company shares profits only with those who "spend the corporation's money, handle its property, make commitments on its account, employ and supervise the employees." To the extent that the system merely increases the incomes of prosperous middle class wage-earners, it loses its appeal from the standpoint of general social improvement.

The second formula is this: *the efficiency of profit sharing varies inversely with the size of the group.* In the absence of visible connection between the effort of the individual worker and the profits of the enterprise, the profit sharing plan relies upon the public opinion of the group to promote efficiency. The slacker is an enemy, not only to his own financial interest, but to that of all other members of the profit sharing scheme. The larger the group, the less alert it will be to detect delinquencies of the individual and the less effective will be its reprobation. For this reason profit sharing among the rank and file of the wage-earners has been effective chiefly in those industries where the labor force is broken up into small, closely coöperating gangs. This accounts in large part for its success in the soap manufacturing concerns like Proctor & Gamble and Lever Brothers, and in construction work, especially in the riveter's gang. An apparent exception to this second formula is found in the case of labor forces where supervision is universally difficult. The Maison Léclaire is a group of house painters and decorators whose men are scattered over the great city of Paris, working largely under their own direction. Similar conditions obtain in the gas companies, where profit sharing by the rank and file has proved effective in reducing costs.[1]

Limitations as an agency of industrial peace. The second major

[1] See, A. W. Burritt, *et al.* *Profit Sharing, Its Principles and Practice*, 1918.

purpose of modern profit sharing plans is to solve the problem of industrial conflict by removing the clash of interest between employer and worker. The logic of the system on this point is clear; it creates a situation in which neither party to the wage contract can prosper without sharing his prosperity with the other. There is no doubt that profit sharing accomplishes this purpose as long as the enterprise flourishes and there are profits to divide. There are few workers whose contentment is not increased by bread and butter benefits. But the great weakness, from this point of view, is that the system seldom survives a period of business adversity and that its breakdown is likely to leave the relations of employer and worker more strained than if no such experiment had ever been tried.

The employees of a given concern, let us say, have been induced to enter into a profit sharing arrangement on the promise that it will pay them dividends in dollars and cents. They respond heartily to the lure of profits and work more faithfully and industriously than before. They are now in the frame of mind which will make them discontented if they receive no reward in addition to their wages; as they view the matter, they have earned a share in the profits and are entitled to it. Suppose now that the business has suffered reverses through some market change which was uncontrollable, or through some error of management. Who would envy the employer his task of explaining this situation to the workers in terms which would arouse their sympathy? If the reverse comes at the end of a long period of success and if it is not too frequently repeated, the profit sharing scheme may weather the storm. Otherwise it is likely to break down, leaving an aftermath of bitterness. Profit sharing is likely to multiply the points of friction between employer and worker. Even if we can assume that it is adopted universally and that profits are generally forthcoming, there is no assurance that conflict between the two parties over the division of the corporate income would disappear. It is just as difficult to devise an obviously fair basis for the division of profits as it is to set up a determinant of fair wages.

One special reason why profit sharing has but limited power to

prevent industrial conflict is that trade-unions are almost unanimous in their condemnation of the system. Their principal official argument against it is that it is merely a device for lowering the normal wage rate. The employer, they say, will never pay labor more than it earns; profit sharing schemes permit the employer to withhold a portion of these just earnings and to assume a pose of virtue when delivering this portion to the worker under the guise of "profits." The result will be a weakening of labor's insistence on a wage adequate to its economic worth. Behind these official pronouncements lies organized labor's real reason for opposing profit sharing, the fear that a spread of this system will disintegrate the labor movement. Employers are largely responsible for this attitude. Profit sharing has often been used, frankly or covertly, for the purpose of smashing the union within an industry. It is safe to say that it will never succeed in destroying the labor movement and that it cannot hope to improve industrial relations as long as it is used for this purpose. If it is to have wide success, it must be one of the fundamental principles of profit sharing that the workers retain their freedom to join trade-unions without prejudice to their standing as claimants of the profits of the enterprise.

Co-partnership. We have seen that profit sharing plans usually have the effect of breaking down clear lines of demarcation between the propertied and wage-earning classes by the process of making capitalists of the laborers. If this process should become sufficiently far reaching to affect a large fraction of the working class, it would nullify one of the principal effects of the Industrial Revolution and automatically remove a potent cause of class conflict in modern society. To this development we may give the name *co-partnership.*

Co-partnership is not merely an aspect of profit sharing. The policy of encouraging their employees to become part owners of the enterprise has been adopted by many large businesses in this country which do not practice profit sharing. Needless to say, this policy is not merely a device to increase the financial power of the business. Most stock acquisition or co-partnership plans

involve special inducements to investment by the wage-earners, which make this method of obtaining new capital an expensive one. Enterprisers can obtain new capital at less cost through ordinary banking channels. The policy is intended to make returns indirectly through its effect upon the attitude of the employee toward the management and toward the business as a corporate enterprise. As with modern profit sharing plans, the chief cause of the rapid development of co-partnership has been the menace of a powerful and aggressive trade-unionism. From the point of view of its broadest social implications trade-unionism is a movement to deprive the enterpriser of his monopoly of the control of production. Its method is bargaining, pressure, force. Co-partnership is a movement in the same direction, but by different means. Through it the workers will gain a measure of control over industry by the same process of saving and risk bearing which has built up the power of the present group of enterprisers.

Examples of co-partnership. A simple plan of stock acquisition leading to co-partnership is that of the United States Steel Corporation. In 1903 this company set aside $2,000,000 of its seven per cent preferred stock to be offered for sale at par to its 168,000 employees. The employees were classified into six groups according to their annual wages. Members of each group were allowed to apply for an amount of stock up to a certain percentage of their wages. These limits varied inversely with the wage income of the group, rising from five per cent for the highest salaried employees to twenty per cent for the lowest paid unskilled laborers. Provision was also made in case the stock was oversubscribed, to give preference in the allotment to the lower classes. Payments were to be made in monthly instalments, not to exceed twenty-five per cent of the month's wages. The company credited the purchaser's account with dividends from the stock and charged him five per cent interest on the unpaid balance. Payment in full was required within three years. Provision was made for refunding money paid on account in case the workman was obliged to discontinue his subscription or left the employ of the Corporation.

The company has made no effort to sell stock to its employees

at unusually favorable prices or to aid them in paying off balances due on account of stock purchases. Since the beginning its policy has been to charge the full market value for each new block of stock offered the employees. Its plan, however, did provide that if the employee retained the stock, remained continuously on the company's pay roll, and showed " a proper interest in its welfare and progress," he should receive an extra dividend of five dollars per share for each of the five years from the date of purchase. Furthermore that part of these payments which would have gone to those who had dropped out of the plan within five years, was also divided among the survivors, " if found deserving," in proportion to their holdings of stock. The details of this plan were altered in succeeding years, but the scheme as a whole has been continued down to the present time. In 1924, 154,342 shares of preferred stock and 463,742 shares of common stock, with a total market value in excess of $40,000,000, were in the hands of the employees of the United States Steel Corporation.

Most concerns instituting systems of co-partnership have offered more generous terms than the United States Steel Corporation. The Standard Oil Company of New Jersey for example has offered to add fifty cents to each dollar invested by the employee. Its plan includes features of partial payment and the receipt of accruing dividends by the subscriber (the normal dividend in this concern is now five per cent) similar to those just described. By 1925, after five years' operation of its stock acquisition plan, about $15,000,000 worth of this company's stock had been bought by its employees. The plans of other concerns contain similarly attractive features as special inducements to investment by the wage-earner. All attempt to make the process of saving as nearly automatic and painless as possible by applying the principle of " stoppage at the source."

Progress of co-partnership. The following statements of the results achieved by some of our largest enterprises will serve to illustrate the extent of the movement toward co-partnership. The figures refer to the situation at the beginning of 1924. In Swift and Company, 21,000 employees — more than one third of

the total pay roll — held shares with a par value of $21,000,000. The American Telephone and Telegraph Company had 94,000 employees making payments on stock, 46,000 of whom held shares paid for in full. In the Du Pont Company $20,765,000 worth of shares were held by employees. Seventy per cent of the employees of the International Harvester Company were stockholders with a total investment of $5,000,000. Practically every workman in the Proctor & Gamble Company was a stockholder as a result of the profit sharing scheme. Over 130,000 shares in Sears Roebuck & Company were owned by the employees either through the profit sharing fund or by direct purchase outside the fund. Fifteen per cent of the common stock of the Philadelphia Rapid Transit Company was owned by the employees. Other large concerns partly owned by the workers were the General Electric, Westinghouse Electric, Eastman Kodak, J. B. Stetson, National Lead, National Biscuit, Republic Iron & Steel, and Studebaker.

Because of its rapid spread and far reaching social implications, the co-partnership movement has been hailed by many writers as "a second industrial revolution." No such verdict is justified on the basis of our present knowledge. The figures given are impressive, but they do not tell us what proportion of the employee stockholders are in the salaried or higher wage classes. No great social significance attaches to a policy which merely increases the possessions of these middle class people. Co-partnership, like profit sharing, is important to the extent that it reaches the typical industrial wage-earner — the proletarian. There is evidence that it is doing this, though not to the extent indicated by the figures given above.

Weaknesses of co-partnership. This movement toward increasing the number of wage-earners who are at the same time capitalists is not without its dangers. Incidentally, since the dominant motive behind most of these stock selling plans is not financial, they may lead to over-capitalization or else to the investment of new capital in the enterprise in excess of the opportunity for its productive use.

But more important than this is the possible effect of such plans

on the economic welfare of the worker. The average wage-earner is in no position to assume risk of loss of his slender savings. Investment in industrial stocks, subject to inevitable risks of fluctuation in value, is unwise for any man whose small savings constitute his only safeguard against a "rainy day." A more appropriate form of investment for the wage-earner would be bonds. But corporations cannot increase their bonds outstanding without multiplying their fixed charges, thus enhancing the danger of turning the co-partnership plan into a disastrous financial venture. Furthermore the essential purpose of the plan is to share control with the workers, and this can be done only through the medium of stock.

The financial weakness of a stock purchase plan from the wage-earner's viewpoint was illustrated forcibly during the crisis of 1920. The breakdown of the market came at a time when large numbers of workers were involved in stock purchase schemes which had absorbed all their savings and obligated them to make payments on account in the future. The depression resulted in wholesale unemployment, which cut off the wages of many of these workers. The partially paid up stock, which represented their only savings, could not be sold in a falling stock market without great loss. Many companies, which in ordinary times would have protected their employees against such losses, were too concerned with the vital problem of saving themselves from insolvency to come to their rescue.

A second objection to co-partnership of this kind contains a fundamental attack upon the purpose of the plan. These arrangements bind the wage-earner too closely to the specific enterprise, thus reducing his mobility and his bargaining power. They are designed in part to reduce the turnover of labor; this is one purpose of their deferred payment features and their provisions of special benefits to those who remain continuously in the company's employ. Within limits this is advantageous to all, for labor turnover is a wasteful practice. But it may be carried to a point injurious to the wage-earner. Having an investment at stake, he is less likely to make the most of his opportunities to improve his

economic position by finding a job elsewhere or by bargaining to raise his wages in his present job. There is danger that he will remain attached to an industry whose fortunes are on the decline. Moreover the plan may be used intentionally by the employer to weaken the resistance of the wage-earner to unjust conditions of employment. If it results in protecting the employer from competition of other industries for the services of his employees, just treatment of the wage-earner will depend too much upon the employer's high-mindedness and sense of fair play. Bargaining between individuals, one of whom has some hold upon the other, rarely results in justice.

This danger will be lessened if the workers continue to be members of the trade-unions. But these schemes are usually destructive of this relationship within the wage-earning class, often intentionally so. They are sometimes so constructed as to penalize the strike and thus further to weaken the resistance of the workers. Such is obviously the purpose of that provision in the plan of the United States Steel Corporation which promises special favors to those who remain " continuously " at work and are " considered worthy " by the management.

Judgment of profit sharing and co-partnership. The fact is that, in our present system of industry, no reliable substitute has been found for the worker's own bargaining power as a means of protecting his interests. Profit sharing and co-partnership plans will not work to the long run benefit of the workers if they undermine labor's self-reliance and self-assertiveness. Both devices will do much to promote industrial harmony and efficiency if used sincerely for these purposes and if they do not interfere with the wage-earner's freedom to combine with others to increase his bargaining power or to strike as a last resort in self-defense. To the extent that they encroach on these powers of labor they will fail of their other purposes by ultimately arousing insurmountable opposition.

Coöperation. The third type of reform mentioned at the beginning of this chapter is much more complex than the two we have discussed. The coöperative movement has taken on such a

variety of forms, has spread so widely throughout the world, and has grown to such proportions that it ranks among the most important social movements of our time. It is difficult to frame a definition of coöperation which will be descriptive of all its aspects. The guiding principle underlying the whole movement however is to do away with the profit-maker in various departments of economic life. The method is so to organize those who are subject to the control of the profit-maker or who rely upon him for certain services as to enable them to take over control of their own economic activities and perform these services for themselves. A complete study of this great movement would fill many volumes. We shall therefore limit ourselves to a brief summary of the methods and results of three important phases of the movement: producers' coöperation, credit coöperation, and consumers' coöperation.

Producers' coöperation. As stated in connection with our discussion of profit sharing, producers' coöperation is a type of industrial organization in which the workers are at the same time complete or controlling owners of the enterprise. It destroys the wage system entirely by making the workers their own employers and managers. It necessarily destroys the profit income as a separate share of the product by merging profits with wages, and it may cause a similar fusion with wages of the other types of income — interest and rent. The enterpriser disappears, and his functions are assumed by the wage-earners. Producers' coöperation is therefore a type of reform favored by those who condemn the wage system as necessarily unjust to the propertyless workers and despair of any improvement short of complete abolition.

Form of organization. The method of organizing a coöperative productive unit is simple. A number of workers get together some capital, buy or rent some land, and set about producing a commodity under a system of democratic control. Since it can be only by accident that all members of the organization invest exactly the same amount of capital, provision is usually made for the payment of a nominal rate of interest to the workers in their capacity as owners. But ownership carries with it no power. The govern-

ing body in the enterprise is elected by the workers, each worker having one vote regardless of his share of ownership or his productive function. By this democratic method, managers are appointed and the system becomes one of self-government in industry.

Experience. Enthusiasm for producers' coöperation as a method of social reform has run in cycles during the past century. The first authentic case on record appeared in England in 1777, when a group of tailors on strike set up a coöperative establishment. The first extensive trial of the system however was made in France during the third and fourth decades of the nineteenth century and especially after the revolutionary outburst of 1848. At about the same time the movement made its first real appearance in the United States. In 1833 the cabinet makers of Philadelphia organized a coöperative association, another was formed among the molders of Cincinnati in 1848, and another among the tailors of Boston in the following year. This period also marked the beginning of authentic producers' coöperation in England. The first number of the *Christian Socialist*, published in November 1850, recorded a needle-women's association, a printers' association, a bakers' association, two associations of builders and two of shoemakers in London alone. In no country did one of these early experiments succeed.

In the eighties a new wave of enthusiasm for coöperation arose in the United States during the ascendancy of the Knights of Labor. Many coöperative associations were founded, but again, with one possible exception, they all either failed or survived only at sacrifice of their essential principle. Many self-governing industries were established in England at this time. Of these, some eighty were in existence and apparently successful in 1912. But none of these have retained the character of true producers' coöperation, since they all allow outside investors a share in the control of management.

In our own day interest in producers' coöperation has again revived. The Guild Socialist movement in England, with its interesting development of the building guilds, has given new

leadership to the movement there. Before the rise of Fascism in Italy, associations of workers for production had begun to make their appearance in large numbers. In America, producers' cooperation is an accepted policy of some of the clothing unions and will probably be the outgrowth of the entrance of the more conservative unions into the field of investment banking.

Cause of failure. As has been said, the record of producers' coöperation is one of almost continuous failure. The reasons have been much the same in all periods of its history. The wage-earners have no adequate appreciation of the difficulties of management and of the rarity of the qualities which make a manager successful. They are loath to pay a wage high enough to attract efficient men to this service. Their own natural leaders leave them because private employers outbid the coöperative association. This same jealousy of the superior man appears in their treatment of the different grades of labor; the scale of wages makes no adequate allowance for the greater importance of the skilled worker. Above all, democracy has always shown itself distrustful of concentrated authority and resentful of discipline. Election of managers by those subject to management has always paralyzed the disciplinary mechanism of any organization. These defects, it will be noted, go to the root of the system of producers' coöperation. If the experiments had failed merely through the wage-earner's ignorance of the detailed problems of producing and marketing a product, the fault might be remediable. But the weaknesses of this form of industrial organization are essential and fundamental.

Reversion to the wage system. The few examples of producers' coöperation which have survived merely emphasize the inherent faults of the system. Prosperous coöperative associations almost invariably refuse to admit new members on a plane of equality and so develop into joint stock concerns employing wage labor. "A coöperative association was formed in 1902 by the employees in the polishing and plating department of the Eastman Kodak Company. This association was placed in the hands of nine directors. Any person wishing to withdraw must first offer his

interest to the company. In less than three years, the number of coöperators was reduced to five; but these five coöperators were hiring men to help them." [1] Such has been the evolution of many other ventures of this kind. This is not true coöperation.

Successful experiments which have retained their original purpose are to be found almost exclusively in lines of industry whose operations are simple, where the skill of the workers is fairly uniform, and where serious problems of management do not arise. It may be well to point out that most of the undertakings commonly called " producers' coöperatives " in this country are not examples of true producers' coöperation. Associations of farmers for coöperative marketing or purchasing, coöperative associations of fruit growers, coöperative cotton gins, creameries, and the like are associations of individual profit-seeking enterprises. They may be effective in increasing the profits of their members, but they exert no tendency to modify the wage system.

Credit coöperation. When we leave the subject of producers' coöperation and follow the development of similar movements in other specialized fields of economic activity, we leave behind us the only phase of the movement which implies far reaching social reform. Other forms of coöperation appeal to middle class sentiments of thrift and desire for accumulation. They involve no essential change in the economic institutions of modern society, but merely an alteration in forms of business organization so that the middleman — the typical profit-maker — may be dispensed with. Coöperation in the business of credit and banking is of this conservative nature.

Schulze-Delitsch and Raiffeisen banks. The model upon which the credit coöperatives of the world have been patterned was supplied by Germany. There the movement is known under two names, the Schulze-Delitsch and the Raiffeisen systems, each taking the name of its founder. The Schulze-Delitsch system embraces the small shopkeepers and business men of the urban districts. The Raiffeisen banks are rural banks supplying credit to small landowners or tenant farmers. This, along with some

[1] Carleton, *History and Problems of Organized Labor*, 1920, p. 242.

slight differences in their business methods caused by the diverse credit needs of rural and urban enterprise, is about the only basis of distinction between them. Each bank is an association of men in need of credit; it lends its credit only to its own members. These members supply the initial capital and make themselves individually and severally liable without limit for the debts of the association. On the strength of this nucleus of capital and the unlimited liability of the members, the association borrows money from ordinary banking institutions or other investors. For obvious reasons it is able to obtain larger quantities of borrowed funds and at a lower rate of interest than its members could obtain if they acted individually. These funds it lends to its own members at a rate of interest sufficiently above the rate which it had to pay to leave a margin for its own costs of operation. Any "profit" is distributed among the members in proportion to their borrowings. The association is self-governing, in the sense that its clients hold control over its policies. The relatively simple problems of administration are cared for by a paid manager, who devotes a part or the whole of his time to this duty according to the scale of the bank's operations.

The bank tends to become a great community centre exerting a strong moral influence over its members and affording a point of contact between them in other than the strictly economic interests of their lives. This is especially true of the Raiffeisen banks, which are located in small country districts where other community organizations do not exist. These banks apply moral as well as economic tests to candidates for membership; no man whose private life is not virtuous and industrious has a chance of obtaining anything from a country bank. But aside from their moral influence, the net effect of the activities of the credit coöperatives is to afford more ample and cheaper supplies of credit to small men of business and thus to reinforce their position in the competitive world.

In Germany and other countries in which similar coöperative institutions have grown up the local credit associations are loosely joined together into nation wide organizations. The larger unit

of the system supplies a forum in which to discuss problems of general concern and to shape policies favoring the spread of the system.

Spread of the system. The growth of the system has been rapid. The Schulze-Delitsch and Raiffeisen banks came into existence about the middle of the nineteenth century. In 1911 there were upwards of 18,000 local branches of these organizations in Germany, all but one thousand being of the Raiffeisen type. The Raiffeisen banks had a combined capital in this year of $650,000,000 and deposits of $570,000,000. The Schulze-Delitsch banks, though much fewer in number, had grown to a larger scale individually. Their total loans in 1911 were a billion dollars, or an average of a million dollars for each society. Since their origin no member or creditor has suffered loss through these banks. From Germany the system has spread to all other countries of Europe and into the Far East. In 1918 Russia had some 20,000 credit coöperatives discharging a vital function in the rural life of the country. The Scandinavian countries, the Balkan States, India, and Japan have rapidly growing movements. Coöperative credit associations flourish in any country where two conditions are present: first, the existence of numerous small scale enterprises in urban industry or widespread peasant proprietorship in agriculture, and secondly, a retarded development of commercial banking affording an elastic deposit currency. In English-speaking countries the efficiency and adaptability of the private banking structure has removed, to a large extent, the need for coöperative credit.

Consumers' coöperation. It is in the field of retail distribution that the coöperative movement has made the most impressive record of success. Indeed to speak of this phase of the movement — consumers' coöperation it is called — as a mere "success" scarcely does justice to the amazing rapidity with which this type of coöperation has spread over the world and gives but a mild indication of the prodigious scope of its operations. In point of the number of people affected, the scale and variety of its undertakings, and the momentum with which it is progressing, this movement must rank among the most important economic

phenomena of all time. Yet its essential principle is simple and its immediate objective modest. Consumers' coöperation aims to displace the retail merchant by transferring his function to the buyers of goods, organized for coöperation. Its object is to save the consumer the profit which the merchant ordinarily receives and thus to permit a slight reduction in his cost of living. The broader undertakings of the consumers' coöperatives have been assumed as an adjunct to this simple function.

The Rochdale system. Great Britain is the birthplace of the consumers' coöperative movement and its present leader. From 1830 to 1840 many ventures in this field of coöperation were begun, but all collapsed because of their failure to understand the economic significance of the merchant's function and to hit upon a practicable form of organization. The success of the movement began with the establishment in 1844 of "The Rochdale Society of Equitable Pioneers." This was a group of twenty-eight indigent working men with a vision of great things to be accomplished through unity and coöperation. So great was their poverty that their initial contributions to the capital of the undertaking averaged only four cents a week apiece. By this laborious process they succeeded, after a year of effort, in gathering together twenty-eight pounds with which to equip a store on a back street in Rochdale. At first they dealt in four staple commodities, flour, butter, sugar, and oatmeal. The store was open for business only on Saturday and Monday evenings, the members serving in turn as salesmen. The following year tea and tobacco were added to their list of commodities, a year later meat. The membership slowly grew until in 1850 it totalled 600. At this point similar associations began to appear in other parts of England copying the structure and the name of the original unit. From these small beginnings the movement has spread throughout the entire world.

Principles of organization and management. The Rochdale system embodies certain clearly defined principles of organization and management, among which the following are of chief importance.

(1) *Capital.* The capital is divided into shares of small denomi-

nation easily within the reach of the poorest working man. In England the par value of the share is one pound sterling. Each member is required to buy at least one share and permitted to buy more. Shares may be paid for on the instalment plan or even by the automatic accumulation of dividends. A modest rate of return — four or five per cent — is guaranteed on this investment.

(2) *Control.* The policy and management of the association is controlled by the members on the principle of one-man-one-vote. This democratic government prevents a lodging of control in the hands of the larger shareholders.

(3) *Business policy.* Normal market price is charged for all goods sold, and sales are invariably for cash. Any saving resulting from the superior efficiency of the coöperative principle as compared with private retail merchandising is allowed to accrue and the proceeds are distributed at the end of the accounting period. Many ventures in this field have attempted to anticipate this saving and reflect it in a cut-rate price policy. These have almost invariably failed. The success of the Rochdale system has been largely due to its conservative business-like price policy and its avoidance of loss through sales on credit.

(4) *Distribution of profits.* The profits of the coöperative stores are distributed among the members in proportion to their purchases, not their investment of capital. With each purchase the customer is given a receipt whose total at the end of the quarter or other accounting period represents his claim to profits. Non-members are encouraged to trade at the stores and receive dividends on their purchases but usually at a lower rate. As stated above, these dividends may be left on deposit with the store as payments toward the purchase of stock. Thus the system becomes one of automatic and painless saving. The dividends average about ten per cent per annum over a long run period.

These principles are varied in detail in the coöperative systems of different countries, but the basic rules of full market price, dividends on purchases, guaranteed interest on investment, and democratic control are to be found everywhere.

Extent of the movement. We can give only a meagre indication

of the spread of consumers' coöperation in England and other countries. At the end of 1920 over 1,500 societies were affiliated with the Coöperative Union of Great Britain. The total membership was 6,400,000, embracing with their families approximately half of the population. The total capital amounted to over half a billion dollars; sales in the preceding year had totaled more than two billion dollars, and the resulting profit had been $130,000,000. Between 1861 and 1914 the net sales of the coöperative stores amounted to over thirteen billions of dollars, and almost a billion and a half dollars of profits had been returned to the consumers. In 1920 the system as a whole employed 200,000 persons and had a wage bill of $130,000,000. In 1914 Professor Gide tabulated over thirty thousand consumers' coöperative societies in Europe alone, with a total membership of ten million people.[1] The movement throughout Europe has doubled its membership in every decade during the past thirty years. In many countries it contains a majority of the total population.

Spread into allied fields. The consumers' coöperative movement of Great Britain has spread out into the fields of wholesale distribution, of production and transportation, of insurance and banking. The Consumers' Wholesale Society of England "has creameries in Ireland, tallow and oil factories in Australia, bacon factories in Denmark and Ireland, great tea plantations in Ceylon and southern India, fruit farms at various points in England, foreign buying depots in New York, in Copenhagen, Aarhus, Odense, Esbjerg and Herning in Denmark, in Gothernberg, Sweden, and in Montreal, Canada. It has a great concession of three hundred square miles in West Africa where it secures the palm oil needed in its soap works. Its buyers are in direct contact with the Greek growers and driers of fruit. It has at Denia, in the centre of the Spanish raisin district, a packing house which employs six hundred persons in picking, packing, and shipping fruit. It owns and operates the largest flour mills and the largest boot and shoe factories and textile mill in Great Britain. It is the largest single buyer of Canadian wheat and has recently purchased ten thousand

[1] Charles Gide, *Les sociétés co-opératives de consommation*, 1917.

acres of wheat land in Saskatchewan. It has its own ships operating between England and France. Each decade it tends to become more self-contained and to produce more and more of what it needs to supply the coöperators." [1] The Banking Department did a total business amounting to $1,347,000,000 in 1916. The Insurance Department carried over 200,000 policy holders in that year. The Scottish Coöperative Wholesale and the wholesale departments of the consumers' coöperatives of other countries do a similarly diverse business, though on a smaller scale than that of the Consumers' Wholesale of England.

Conservatism of consumers' coöperation. How can we account for the amazing success of this great movement? In the first place it must be said that the success of the movement has been due primarily to its conservative practical policies and not to any far-reaching program of social reform which it contains. The largest single element in its success, at least in England, has been its appeal to the thoroughly middle class desire to accumulate capital and to its provision of an almost automatic method of saving. The system has demonstrated beyond all dispute its ability to lower the living costs of the average English family by at least ten per cent. So simple and easy is the process of saving this portion of his normal expenditure and allowing it to grow at compound interest, that the working class member of the system becomes a small capitalist almost without his knowledge.

The conservative nature of consumers' coöperation is further shown in the management of its larger productive undertakings. The system is not one of true producers' coöperation. In its factories and on its farms, plantations, and ships, the Consumers' Wholesale employs ordinary wage labor which is outside its membership. It deals with these workers much as would any other capitalist employer, hiring labor on the basis of efficiency, exercising its right of discharge, and insisting on discipline. It has never adopted a system of profit sharing in these undertakings. Though it has shown more sympathy with the aspirations of the trade-union than the ordinary employer and has set the pace in its wage

[1] Harris, *Co-operation*, 1920, p. 223.

policy, the Wholesale retains in its own hands the rights of managerial control. It has even had strikes in some of its establishments, and it cannot be said to have worked out as yet a policy of industrial relations which is entirely acceptable to organized labor.

Avoiding the wastes of competition. In the second place, the movement has brought to light certain wastes inherent in our present system of marketing and has shown the superiority of the coöperative principle in this field of economic activity. It must be said however that the negative results of coöperation in marketing are almost as significant as its positive gains. The experience of the coöperators has disproved the common opinion that the ordinary marketing structure is costly because of profiteering practices by the merchants. Their gains have come about, not through a squeezing out of the merchant's profit, but by saving wastes inherent in the competitive system which private merchants cannot avoid. As all students of business economics are well aware, the marketing side of our productive system has responded less than any other to the pressure for greater efficiency. So far as the costs of marketing can be traced to the wastes of competition, they can be reduced by consumers' coöperation. There are some types of cost which the coöperators can avoid. Since the customers of the store control its buying, they know in advance what kinds of goods will be in demand and in what quantities and can adopt the least expensive means of getting these things, avoiding also the loss of unsold stock. They are sure of their patronage and do not have to practice expensive policies of advertising and salesmanship to attract custom from other merchants. Advertising and selling devices, when used by consumers to facilitate the satisfaction of their wants, have a different function from that of a system of marketing where competing producers contrive to make the consumer want what they have for sale.

The wastefulness of competitive marketing is especially great in the older countries of Europe, where the social stigma attaching to this form of activity tends to repel the able enterprisers and leave the function in the hands of mediocre men. In these coun-

tries the customary practice of extending credit to customers is a source of loss; business methods in retail merchandising are traditional, lethargic, time-consuming; shops are managed by rule of thumb. Accordingly the opportunity of improvement through a system managed with greater alertness and on stricter business principles is especially great. Consumers' coöperation has attained its maximum growth in such countries.

Broader social aspects. What has been said about the conservative and narrowly practical character of consumers' coöperation must not be taken to mean that this movement is not a power among the broad social forces of the day. The movement is distinctively working-class in its origin and present personnel. It is the largest single working class organization in many of the countries in which it has gained a foothold.

Of necessity, its members develop an *esprit de corp* which can be utilized to unite them for collective action along other lines than that of mere retail merchandising. Because of its democratic control, the coöperative movement is especially susceptible to the idealism and the reform tendencies of the rank and file of humanity. Hence we find it involved in many political and social activities which are quite apart from its strict function. In these phases of their policy the coöperatives take their tone from the prevailing temper of their membership. They will be socialistic in countries where the mass of working people are radical, mildly progressive in more conservative countries. These broader activities do not spring from the principle of consumers' coöperation, but merely represent the universal tendency of social institutions to take on new functions as they are made the medium for expressing the many sided interests of their members. Yet so gigantic has this movement become that these supplementary activities must be reckoned among the most powerful of modern social forces.

Social program in England. The coöperatives were leaders in working class education in England. From the beginning they appropriated two and a half per cent of their profits for educational purposes. They now hold scholarships at Oxford University, support liberally extension schools and colleges for working men,

and have recently established a college of their own at Manchester. In the conflict over industrial relations they are on the side of labor. They have loaned funds or supplied food to men on strike, and their aid has been a large factor in promoting the solidarity of labor and increasing its strength during the trying period of readjustment which has followed the World War. Politically the Coöperative Union supports the Labor Party with its program of constitutional socialism. Its influence on current radical tendencies in England was shown in connection with the development of the Building Guilds. These organizations, producers' coöperatives with broad socialistic objectives, were saved from collapse at the outset through the willingness of the Banking and Insurance Departments of the Coöperative Union to finance and underwrite their enterprisers.

In other European countries. The coöperative movements of the other European nations have been swept by the idealism of their members into similarly far-reaching social activities. In Belgium the profits of the coöperatives are not distributed but invested in the Maisons du Peuple. These are great community centres, supplying the working classes with facilities for amusement, education, and cultural activity. The Maisons give free medical and clinical service, provide relief for men on strike, organize and finance travel parties for young people at home and abroad. The movement in France has strong political affiliations and is recognized in the administrative machinery of government through a Council of Coöperation in the Ministry of Labor. So it is with other countries. The forms taken by these social and political phases of the movement as well as their effects on the established institutions of capitalistic society vary in conformity with the peculiar social conditions of the different countries and with the prevailing temper of the working classes. But everywhere the movement has become a power among the social forces of our time.

Obstacles to consumers' coöperation in the United States. America has remained largely untouched by the movement whose impressive record in other lands has been sketched in the preceding

pages. Until 1900 most American ventures in this field were short-lived; nothing approaching a movement of national scope existed. From 1900 to 1914 the appearance of coöperative retail stores clustering about the chief industrial cities gave the first real evidence of vitality in the movement. Since the outbreak of the World War the spread of the coöperative principle has been fairly rapid and well sustained. The Coöperative League of America in 1920 estimated a total of between 3,000 and 4,000 retail stores, with a combined capital of $15,000,000 and a membership of some 500,000. This is a small record of achievement in comparison with the situation in other countries. It is noteworthy also that the coöperative movement in this country flourishes principally among foreign born wage-earners who have been trained to the system in other countries.

The reasons for this slow growth in the United States are inherent in conditions peculiar to our country. They may be summarized as follows: (1) Relative efficiency of the marketing system. Retail merchandising in our country has been more progressive than in Europe. Merchandising has attracted many of our ablest enterprisers. Department stores, chain stores, and mail order houses leave little room for the saving of cost in distribution. The coöperatives of England are just now beginning to feel and to fear the competition of these forms of efficient profit making enterprises. (2) The great extent of the country has been a handicap to any unified movement among consumers. (3) The greatest obstacle has been the individualistic and impatient character of the American population. We have had no stable wage-earning class with sufficient solidarity to support such a movement. The tendency to migration has continually broken up local community groups. Abundance of economic opportunity has attracted the wage-earners to positions of independence in the economic structure, and private enterprises have lured away the more able labor leaders by offers of higher salaries. The higher standard of living and the greater ease of life have prevented the development of the thrifty, penny-saving habits to which such a movement as this makes a strong appeal.

CHAPTER LVI

SOCIALISM

Breadth of the subject. Socialism is a complex and many-sided movement, including so great a variety of practical and idealistic aspirations and assuming such diverse forms in different places and at different times that it is difficult to define it in a single formula and quite impossible to describe and analyze it fully in the small space at our disposal. It is to be understood only as a strong current in human thought and behavior, a movement or trend involving millions of people and carrying them along different paths of social change. Its rational program is not its most important feature. Its emotional content, the idealism of its followers, its dynamic power, are the features of chief significance from the standpoint of its present and future influence upon society's social organization. Since these features cannot be brought to light by a critical analysis of the economic doctrines and the paper programs of socialist parties, they receive too little attention in most studies of this movement. We too are obliged to leave them out of account in the discussion which follows, since our attention must be focused principally upon the economic aspects. But it is well to begin our study fully aware that no adequate appraisal of this movement can be based on economic analysis alone.

Economic analysis of socialism. Socialism contains an indictment, an analysis, a program of reform, a campaign. Its *indictment* is a list of charges against "any and all industrial systems based on private property and competition." In its *analysis* it traces the development of capitalism back to its origin, dissects its present structure, and stresses certain developmental tendencies which, so it is argued, must inevitably destroy our present order. Its *program* sketches the outlines of a substitute for the capitalistic system which socialists believe will remove the evils of modern

capitalism. Its *campaign* is a plan of immediate tactics for those who adhere to its cause.

In no one of these phases of the movement are all socialist groups in agreement, though unanimity of opinion is much more nearly approached with regard to the first two phases than the last two. Socialists differ, so far as their indictment is concerned, only as regards the inclusiveness of their charges against capitalism and the relative stress they place on its different faults. In our day the analysis is based principally on the interpretation of capitalistic society laid down by Karl Marx; hence we find considerable agreement here.

The programs differ widely in detail; socialists are not at one in their specifications of the ideal society of the future. But if we disregard these differences of detail, we can discern an essential similarity in the various socialist programs. They all propose a type of society in which there will be no private property in wealth used for productive purposes and of course no inheritance of these forms of wealth. Furthermore they agree that the productive system of the future shall not be governed by competition and private initiative, but by some sort of collective control and guidance. Socialism then is destructive of these two basic institutions of our order, private property and free competition or individual enterprise. The programs differ in their substitutes for these institutions. (1) What group shall own the productive wealth of society when private ownership has been forbidden? (2) And through what mechanism shall society decide upon the kinds and quantities of goods to be produced and fix their prices when these things are no longer left to the free play of competition? It is over these questions that the socialist movement breaks up into different schools.

In regard to their campaign socialists divide into two groups: those relying on peaceful, evolutionary methods, and those who place their reliance on violence and force. The former are *constitutionalists*, believing that gradual increase in their power on the political field will eventually bring them to their goal by thoroughly legal means. The latter are *direct actionists*; they have no faith

in politics, being convinced that only through force or threat of force will the evolution come.

Definition. Any definition which attempted to embrace all these phases of the socialist movement would be too cumbersome for practical use. We shall define the movement in terms of its program, stressing the essential similarities between the different socialist schools. *Socialism is a program of reform which repudiates private ownership of the means of production and competitive control of industry; it proposes a type of society in which productive capital will be owned collectively and economic activity will be controlled by authority.*

Socialist criticism of capitalism. The socialists have no quarrel with capital. They believe, however mistakenly, that the supply of capital will be increased and its efficiency improved under their system. Their attack is upon capitalism — the organization of society — not upon the material instruments of production. Like all reform movements, socialism has its "lunatic fringe," whose indiscriminate denunciation of all that is and whose panegyric of the Utopia of the future do not merit serious consideration. There are however many points in the socialist indictment which are thought-provoking to every student of society. The chief of these criticisms are summarized in the following paragraphs.

(1) The competitive system is wasteful of resources and effort. In established concerns individual enterprise results in duplication of effort. Competitive advertising, competitive salesmanship, and other forms of competitive activity increase the costs of each enterprise. These costs could be reduced if the services common to all producers were provided through a common agency. In the competitive system the superior business methods and most efficient machines are controlled by individual enterprises. Increased efficiency would result if the best processes were made the common possession of the whole industry. Thus far the argument is much the same as that which contends for the superior efficiency of combination and large scale production. In a more general sense the socialists urge that the present system is wasteful because in their search for profit enterprisers direct capital and labor into

fields of activity which are already well enough equipped to serve the needs of society. It would be more economical to use collective intelligence in forecasting our needs and conscious coöperation in adjusting the means of production to the ends in view.

(2) The competitive system results in wasteful strife. Employers and workers strive with each other over the division of income from industry. Producer strives with producer for advantage in the market. Out of this strife spring practices which result in loss of productive forces, as we have seen in our study of industrial relations and of governmental control of competitive practices. The anomaly of the situation is that these contesting parties are really coöperating toward common ends. Intentional coöperation is less wasteful than antagonistic coöperation. From a broader viewpoint, socialists believe that the competitive system is the root cause of international war with its incalculable toll of destruction and waste. They believe that their system will automatically produce harmony between individuals, social classes, and nations.

(3) The competitive system is destructive of human and natural resources. This criticism is implied in part in the foregoing statement. But, in a more general sense, it is urged that the pursuit of immediate private gain leads to the exhaustion of irreplaceable natural resources and the wasteful exploitation of human energy, especially that of children and the more impoverished male and female workers. The modern conservation movement and modern labor legislation imply a similar criticism of uncontrolled competition.

(4) The competitive system is unstable and erratic. Socialists argue that the business cycle is the result of distorted production, which in turn results from anarchy in production and inequality in distribution. Seasonal and cyclical fluctuations, with their recurrent losses from idle capital and labor, could be prevented in a system controlled and guided collectively.

(5) Private property results in great inequality. Private property includes the right of inheritance. As the socialists view the

matter, the great inequalities in incomes are caused by private ownership of income-bearing wealth and perpetuated by the institution of inheritance. The socialists believe that capacity is distributed with approximate equality among the different income classes. Opportunities for self-development cannot be equalized so long as some men are born to a position of privilege while the many are born in an environment of poverty, ignorance, and hopelessness. This situation is not only unjust on ethical grounds but it is inefficient and precarious. Inefficient, because it smothers the latent abilities of those in the lower groups and subsidizes idleness in the upper classes. Precarious, because it arouses the envy and resentment of the victims of the injustice, who are an overwhelming majority and have the power to destroy the system.

(6) The wrong things are produced. Inequality of distribution distorts the whole productive system out of alignment with the real needs of society. Only those things are produced which will sell at a profit. Hence the kinds and quantities of goods resulting from the combined activity of mankind are determined by the spending power of individuals and not by the test of what is needed to nourish a strong, healthy population. There are abundant examples of the grotesque effect of mal-distribution on production. A serious housing shortage exists at a time when much labor and capital are being devoted to the construction of steam yachts and country palaces. The average family suffers from a shortage of nourishing food and serviceable clothing while useless or injurious articles of diet and frivolous articles of fashion, are supplied in abundance to the rich. Foolish expenditure of all classes has the same distorting effect on production. The production of chewing gum, cosmetics, and patent medicines absorbs capital and labor which ought to be supplying impoverished people with the means of wholesome living. The same evil is to be seen in the fact that many things of vital importance to human welfare are not produced at all because no individual enterpriser can make a profit out of their production. The socialist system proposes a quite different basis of production. The social group will decide what things are needed to support a wholesome life, and these things

will be produced first in such abundance that there is plenty for all. Only surplus labor and capital will be devoted to the production of luxuries and extravagances, while injurious goods will not be produced at all.

(7) Capitalism is destructive of liberty. The defenders of the present system of industry, so the socialists charge, are accustomed to speak of it as a system founded on individual liberty. It is said that our society permits each human being to work out his destiny according to the powers that lie in him unrestrained by bonds of status and tradition. This the socialists deny. To them the present order means the enslavement of the great majority of propertyless people by the owners of wealth and the subjection of the whole body of society to the dictates of those who dominate its essential industries.

First the wage system is condemned as a species of slavery. It is explained as an arrangement between the few who own the instruments of labor and the resources of the earth and the many whose only hope of survival consists in their getting access to these things. The outcome of the arrangement is that the many deliver themselves into the power of the few, their productive energies to be exploited to the profit of the masters. The freedom of the individual worker is freedom to starve if he does not wish to surrender.

Secondly the economic destiny of the social group as a whole is at the mercy of the more ruthless or more fortunate individuals. The landlord exacts his tribute for the privilege of using the gifts of nature. Those who have acquired or inherited ownership of socially necessary instruments — railroads, steel mills, factories — can work their will upon the rest of mankind. These essential instruments are social in their purpose and importance; they cannot safely be allowed to fall into the hands of individuals.

Finally all social classes — wage-earners, employers, capitalists, landlords, consumers — are seen to be slaves to the passions of self-interest and acquisitiveness which the competitive system arouses. The socialists believe that the order of society which they propose will be one of real freedom. Men will not work for the

profit of other men but for the group as a whole and under its commands. The instruments of labor and the resources of nature will be owned and managed collectively under democratic government. The passions of greed and self-seeking will be held in check by arrangements which play upon better qualities in human nature.

The doctrines of Karl Marx. The socialist analysis of the development and present tendencies of the capitalist is derived from the work of Karl Marx, who was the founder of all modern schools of Socialism. *The Communist Manifesto*, written in collaboration with Frederick Engels in 1848, and *Capital*, an analysis of the economic laws of capitalism and a forecast of its future development which appeared in 1867, have profoundly influenced all subsequent radical movements. For our purposes, the essential doctrines of Karl Marx may be presented under three heads: the materialistic interpretation of history, his system of economics, and the laws of capitalist development.

Materialistic interpretation of history. Karl Marx was a forerunner of the modern group of historians and sociologists who explain societal evolution in terms of economic factors. Throughout all human history the framework of social institutions, according to this doctrine, has received its peculiar character from the way in which the people have made their living. Government, religion, family system, art, literature, and morals are by-products of man's economic activities, harmonizing with the prevailing system of production and exchange. The changes which have appeared in these institutions in the course of societal evolution have been the effects of prior changes in the economic basis of society. The most important events in human history — wars, the rise and fall of empires, the Crusades, the discoveries, and the migrations of peoples — are traceable to economic causes. This doctrine, in some degree of comprehensiveness, is now believed to be true by most students.

The class struggle. But Marx carried the doctrine one step further. To him not only were the economic institutions basic and determining, but these institutions had their own vital centre in the class struggle. Every system of production which mankind

had devised consisted in a hierarchy of classes — exploiters and exploited. In classical society, slave, plebeian, patrician, and noble; in the Middle Ages, serf, vassal, knight, and feudal lord, were ranged in an ascending order of increasing privilege. The interests of these classes clashed and forced them into continuous conflict. Revolutions in economic institutions, themselves the cause of all great changes in other institutions, resulted from this condition of class conflict when a previously exploited class gained power to throw off the yoke of the oppressor. Thus, "the history of all hitherto existing society is the history of class struggle." But each such revolution merely reëstablished the system of exploitation in a new form. The last great change of this type occurred when the bourgeoisie arose and overthrew their feudal masters thus instituting the system of capitalism. Out of this grew the present class alignment — capitalists and laborers — and the present struggle between these two classes. The outcome of this struggle lies in the future, but it is predicted that the wage-earners will eventually succeed, as did the present capitalists, in overthrowing their exploiters. Their triumph will mark the end of the class struggle, since they will re-organize the industrial system so as to remove the possibility of exploitation.

Marxian economics. The base of Marx's system of economics is the labor theory of value. Normally and in the long run, he says, goods are exchanged for each other in ratios which measure the relative amounts of labor power incorporated in them. Various exceptions are acknowledged, and it is admitted that market prices may temporarily depart from this value norm. But on the whole it is labor cost, including in this concept a proportionate part of the cost of providing capital instruments and raw materials used in production, which controls value.

The law of wages is a corollary of this law of value. Labor, like other goods, sells at a price equal to its cost of production. The wage rate is determined by the cost of "the necessaries required for the laborer's maintenance plus another amount required to bring up the quota of children who are to perpetuate the race of laborers." Allowance is made for the standard of living as far as

it affects the birth rate. This is the "iron law" — or subsistence cost theory — of wages.

Exploitation in capitalist society takes the form of surplus value. The wage-earner is hired on a cost of production basis, but the wage contract places him in the power of the capitalist to be used as the capitalist pleases. The worker is able to produce enough to return his cost in a fraction of the working day. But he is kept at work for a longer period, his products being appropriated by the employer. This part of the product therefore is obtained by the capitalist without cost to himself. It is sold, like other products, at a value determined by the amount of labor it contains. Here we have the concept of surplus value. The incomes of interest, rent, and profits are drawn from the sale of this surplus product of the wage slave. The recipients of these incomes are parasites, exploiting, not the consumers as is sometimes believed, but the laboring class.

This whole structure of economic theory is, as will doubtless be recognized by the reader, fallacious. The value theory is unsound, because it leaves out of account the concept of marginal utility. Wages, as we have learned, are not determined by the cost of living but by marginal productivity. Opinions may differ as to whether interest, profits, and rent are justifiable incomes, though we have seen reason to believe that the first two at any rate are necessary payments for essential functions. In any case, the Marxian explanation of their existence and their amount is erroneous. Until very recently socialists have been content to repeat these economic fallacies; now they are beginning to perceive their inaccuracy and to discard them. The strength of modern socialism lies elsewhere than in its system of economic theory.

Laws of capitalist development. The conclusion of Marx's diagnosis of the capitalist order was a conviction that socialism is the inevitable next step in societal evolution. His system made a sharp break with all preceding socialistic programs. Before his day socialism had taken the form of Utopia building, a construction of little communities in which a perfect order of society was to demonstrate its power to evangelize the world. The Utopians

believed that all that was necessary to usher in socialism was suddenly to call a halt on the operation of the present system and to transform it over night into an ideal commonwealth. Their experiments failed, and their propaganda lost its appeal to all but the most visionary. Marx repudiated this whole system of thought with scorn. It is quite as impossible, he said, for the prophet and the evangelist to produce socialism artificially as it is for the alchemist to turn baser metals into gold. Socialism must come as a natural outgrowth of capitalist society or it will not come at all. Socialists may speed up the intermediate stages leading to the new society and may facilitate the transition, but the motive power must be supplied by the forces of capitalism itself. Marx thought that he had discovered the existence of two such forces tending toward socialism — (1) the tendency toward concentration and (2) the increasing number and misery of the laboring class.

Tendency toward concentration. It is argued that capitalist development tends in two directions toward increasing concentration; first the concentration of industry into larger and larger business units; secondly the concentration of wealth and income in the hands of a diminishing number of people. Industrial concentration is said to be due to the superior efficiency of large industries, which enables them to undersell and crush out smaller competitors. This prediction was made before the trend toward large scale production had set in. Marx set no limit to the process; we are to infer that concentration will continue until a single gigantic business unit monopolizes each of the major fields of industry including agriculture. In agriculture the prediction has been falsified by subsequent events. How far it has been verified in other fields we have seen in another place (see especially Chapter VII). This tendency was to lead to socialism for the double reason that collective control would become both necessary and easy. Necessary, in order that society might protect itself from the menace of these monopolies; easy to institute, because of the disappearance of effective competition, and easy to administer because of the relative simplicity of operating the few centralized productive units.

Concentration of wealth was inferred, partly from the preceding tendency, partly from other evidence. Marx made the mistake of assuming that concentration in industry of necessity involved concentration of ownership. He wrote before the day of the corporation, whose joint stock principle permits a growth in the scale of the business unit coincident with an increase in the number of those who hold property rights in it. Great as are existing inequalities in distribution, we have no evidence that they have grown greater since Marx's day or are now growing greater.

Increasing number and misery of the laboring class. The tendency toward concentration coupled with other features of the capitalist system was expected to produce an expanding propertyless class, with a steadily declining economic status. The crushing of small enterprises by the growing monopolies would precipitate former small capitalists into the laboring class. The growth of an industrial reserve army, the effects of periodic depression, and the absorption of an increasing fraction of income by the rich would cause increasing misery for this ever-growing group. The rich would grow richer and fewer; the poor more numerous and poorer. This process would cause the system to break down through sheer inability of the masters "to maintain their slaves in their slavery." It would also cause the victory of the proletariat to put an end to the class struggle, since it would be a victory of "the immense majority in the interests of the immense majority." We need not repeat our proof that this theory of increasing misery has not been verified by the event. Marx did not foresee the growth of trade unionism, social insurance, and labor legislation, which would protect the most defenseless of society's members from the consequences of their own bargaining weakness. Until the opening of the twentieth century the average workman's absolute income increased steadily, and his relative position in distribution at least did not decline. It is yet too early to say whether the developments of our own generation will reverse this process.

The Marxian analysis of capitalism may thus be stamped as false in many particulars. Yet its importance in shaping the development and the present character of socialist thought cannot

be denied. It placed socialism on a new and firmer basis. Its specific predictions may be rejected as not proved, without destroying the value of its central thesis; *i.e.*, that every thorough evolution in society's institutions must come about as a natural unfolding of forces at work in the preceding order. This principle has had great influence on the practical tactics of Marx's followers. It has convinced them that capitalism must run its course before socialism is possible. The factory system in industry with its corollary of a large wage-earning class, democracy in the political field; education for the mass of mankind, experience in collective action gained through labor organizations — these phases of capitalism must be nourished to maturity. Violent premature revolutions will fail disastrously. Marxian socialists have been forced by the logic of their system of thought into peaceful and legal forms of activity.

Socialist programs. We have used the term program to denote the substitute industrial order with which the socialists propose to displace the present capitalist system. The socialist movement of the world divides into well-defined groups or schools over the question of program. However, as shown by our definition of socialism, the programs of these different schools have common elements which bind them all into a single general movement of social reform. They all propose to abolish private ownership in the means of production and to substitute for competition some type of coöperative control over productive processes. They differ as regards the form of collective ownership and the mechanism of control which they propose to set, and these differences permit their classification into certain major groups, which we may now proceed to examine.

Collectivism or state socialism. The largest single group comprises those who would make the present territorial state the legal owner of productive instruments, and would vest control of productive processes in chosen representatives of all the people. This type of socialism is known as collectivism or state socialism. Its program provides for division of ownership between different political units. The central government would own the indus-

tries whose significance is nationwide; to the smaller territorial divisions and the municipalities would belong those which supplied only the local markets. Some allowance is made for individual enterprises, provided they do not employ hired labor and sell their products under strict price control. In this system wages would be the only possible type of income, and everyone who wished to escape starvation would be compelled to work, excepting of course the aged, the children, and the defectives who would be cared for by the state. The state would be the dominant power in the lives of all; it would be the only employer, the owner of all land, the only saver of capital, the only seller of goods. It is intended that this great power shall be brought completely under democratic control through universal suffrage, direct election for short terms of all policy making officials, proportional representation, single chamber legislatures, and other devices.

In saying that collectivism is the largest single group of socialists we have included in this group those who accept the central principle of political ownership and control but modify in some details the more extreme doctrines of state socialism. Many factions within this large group are opportunists in outlook. They press for an extension of government ownership and control wherever private ownership appears to threaten popular welfare and competition seems ineffective or destructive. They have no well-defined general objective and do not concern themselves with the final outcome of the process. On the other hand are those who have in their minds a complete and logical system of social organization, but recognize the limitations of the principle of government ownership. Their aim is to leave as large a sphere for individual initiative as is consistent with the general principle that land and the larger aggregates of capital be used for the " common good " instead of for private profit. The Fabian Society of England is best representative of this school of modified collectivism. This small group of intellectuals, numbering among its members such people of international prominence as Sidney and Beatrice Webb, Philip Snowden, Bernard Shaw, Graham Wallas, and William Clarke, has given much study to practical administrative problems of

socialism. They have supplied the intellectual leadership and much of the program of the British Labor Party and provided members of the Labor cabinet during that party's brief tenure of power in 1924. The majority socialist parties of Germany, Scandinavia, and English speaking countries are collectivist in principle.

Syndicalism. A reaction against the political dictatorship of state socialism is seen in the program of the Syndicalists. This school of socialism is a product of the organized labor movement ("syndicat" is the French name for the trade union). The central principle of the syndicalist program is the ownership and control of each major branch of industry by the organized workers in that industry. The territorial state would be shorn of most of its functions, if it did not entirely disappear, and society would be resolved into large, independent, self-governing industries. The local units of workers in each of these industries would be given a substantial degree of autonomy, though they would be united with other local units into larger territorial and international groupings for the determination of general policies. For coöperation between the different industries, the syndicalists rely upon loose voluntary federations among the several industrial organizations. Syndicalists fear the state and distrust all forms of political action. Their campaign is one of direct action, often violent and destructive of property. Their principal weapons are sabotage and the strike, and for the final transition to socialism they rely on a simultaneous strike of all workers, accompanied by violence and terror. This branch of the socialist movement has its centre in France and the other Latin countries of Europe. It is represented in the United States by the Industrial Workers of the World.

Guild socialism. As a compromise between the collectivists and the syndicalists we find guild socialism. It shares with the syndicalist their distrust of an all powerful state and their reliance on the trade union as the nucleus of future industrial society. But it also recognizes the danger inhering in a social organization in which each of the great industrial services is monopolized by its particular group of workers without effective machinery to enforce coöperation between them. The guild socialists therefore propose

a bilateral organization of society. The democratic state is to remain as a protector of the interests of the general consuming public and a supplier of such national services as education, police, and foreign relations. On the other hand, each major branch of industry would be managed by a union of all its workers, manual and intellectual, operating in conjunction with the state, which is to be legal owner. Local units would be united into national guilds, and the different guilds would be drawn together for coöperative action into a Guild Congress representing all branches of the economic activity of society. The Guild Congress and the political Parliament would counterbalance and check each other; contact between them would be furnished through standing joint committees. This, in bare outline, is the system proposed by the guild socialists.

Guild socialism is of late development. It arose as a distinct school of socialist thought in England during and after the World War, when a small group of brilliant university men gave a strong initial impetus to the spread of its doctrines through their writings and public addresses. The system makes a strong appeal to existing labor organizations with a socialistic bent which shun the violence and anarchy implied in syndicalism. One interesting feature of the movement has been the organization of self-governing guilds of workers in the building trades of England. This new experiment in producers' coöperation began during the reconstruction period following the war, as a means of supplying the shortage of houses. The building guilds for a time demonstrated their ability to underbid private contractors in the construction of dwellings, while yet affording continuous pay to their members during periods of idleness caused by the seasonality of the industry. Their early success attracted workers in other lines, especially the related industries producing household equipment and building materials, into similar ventures. The movement however declined rapidly with the increasing severity of business depression through which England has been passing. In America the powerful railroad unions favor a scheme of organization for the transportation system — the so-called Plumb plan — which is guild socialist in principle, and the

United Mine Workers are pressing for a somewhat similar plan of organization in the mining industry.

Communism. The dramatic triumph of revolutionary socialism in Russia has attracted world-wide attention to a fourth type of socialist program — communism. Before the revolution the Bolshevists, who called themselves followers of Marx, had a program of modified state socialism. They differed from other Marxian groups in their repudiation of the doctrine that socialism could be established successfully only in countries with mature capitalist institutions, this condition being obviously not fulfilled by Russia, a semi-feudal country. They believed a successful revolution was possible if carried through by force and maintained by the dictatorship of a determined, well-disciplined party. This group took the name Communist Party after their successful *coup d'état* in 1917. The system of doctrine which we now call communism was developed by a process of adjustment to practical problems arising in connection with their attempt to govern the country.

If we may safely generalize from the policies of the Communist Party during the height of its power in Russia, communism in its purity implies the following type of social organization. Private property is abolished, in consumption goods as well as in capital and land. The state is owner and operator of every form of industrial enterprise and the sole distributor of products. Labor gives the sole right to income, and all labor is performed for the state. Wages are to be the same for all workers, regardless of function, and are to be paid in rations of goods or in labor tickets which give the holder right to demand goods at the public storehouses. Money is to disappear. The only remaining form of exchange is that of "service for counter-service"; that is, labor service for the state by the individual and counter service in the supplying of consumption goods to the individual by the state.

Failure of Russian communism. From 1918 to 1920 the Russian communists employed measures of increasing severity in their effort to force this system upon the country and make it function. In the end they failed, because of their inability to compel the peasants — an overwhelming majority of the whole

population — to supply farm produce to the state at nominal prices. Confronted with famine in 1920 and with a critical shortage of raw materials in all urban industries, the communists hastily surrendered their program. Since that date, the economic life of Russia has been organized under a system of qualified state socialism. The state retains ownership of the more essential industries, but a free market has been restored; private enterprise is reëstablished in the uncontrolled industries, especially in agriculture; the wage system is reëstablished under a stringent labor code, which guarantees high minimum standards to the workers.

The soviet organization. Because of this change of front and the doubt it casts upon the theoretical consistency of the communist program, this school has come to signify chiefly a type of social structure and a method of revolution. In structural organization, the unit is the *soviet*, a committee of workers representing the different economic functions of the local community. On the basis of these local soviets is erected a hierarchy of federal units, district and regional, and finally the sovereign national Congress, possessing legislative, judicial, and executive powers. Between sessions of the Congress, these sovereign powers are delegated to a Central Executive Committee, elected by the Congress from among its members. The Central Executive Committee is divided into sections, each having charge of one of the administrative departments of the state, and the Chairmen of these sections constitute the powerful Council of People's Commissars, which is the cabinet of Russia. The places on this Council have been held by the leaders of the Russian revolution, Lenin, Trotsky, Tchicherin, Krassin, Stalin, and others. Theoretically the Council is responsible to the Central Executive Committee and through it to the Congress. Actually it has had virtually dictatorial power over Russia.

The method of communism. Revolution by direct action is the method of communism. A small, determined, ruthless party is to seize power in the name of the proletariat, whose dictatorship over all social classes they are to establish. The dictatorship is to consist in reality of the supremacy of a nucleus of self-appointed leaders, for the rank and file of the party is trained to follow these leaders

with implicit unquestioning obedience. The dictatorship is to continue until capitalist institutions have been transformed and the people converted to the doctrines of communism, when in theory it is to give way to a completely democratic system of government on the soviet model.

Extent of the socialist movement. There is no industrial country in the world in which the socialist movement is not represented. In every country also the socialists are divided into factions whose relations with each other are anything but peaceful. In Russia the Communists, by their own count, number not more than three quarters of a million in a population of some one hundred and twenty millions. This small fraction of the people maintains itself in power, partly by the use of terroristic methods, but chiefly because of the apathy of the peasants, who constitute the bulk of the population. Other socialist parties have been under the ban but are now beginning to reappear as the Communists relax their dictatorship.

In Germany. The Social Democratic Party, a Marxian group organized on a program of collectivism, is the majority wing of the German socialist movement. In the last election before the World War, the Social Democrats polled the largest popular vote of all political parties. This party reached the climax of its power in 1918, when it directed the republican revolution, and during the succeeding months, when it was virtually in command of the country. The new government adopted a mildly socialistic program in domestic affairs, but the policies of socialization were not completely carried through, because of the breakdown of the government under the insuperable difficulties created by the peace treaty. In subsequent elections the Social Democrats lost their dominance as a political party, though they still hold the balance of power among the numerous groups in the Reichstag. The socialist movement had split in 1915 over the war program of the government, a more extreme party withdrawing from the main group under the name Independent Socialists. This minority faction has since become a Communist Party, in close affiliation with the Communists at Moscow. In the presidential election of April,

1925, the Social Democrats received a popular vote of almost eight millions and the Communists a vote of about two millions, out of a total of twenty-seven millions.

In Central Europe. The socialist movement has become a powerful influence in political affairs throughout Central Europe. During the closing months of the war, temporarily successful communist revolutions were carried through in certain countries most seriously affected by the prevailing chaos. These soviet experiments failed or were suppressed by force, but the more conservative socialist parties emerged from the period of reconstruction occupying a position strong enough to affect the policies of all the new governments. In Austria the different socialist groups, led by the conservative Christian Socialists, are in the majority. A socialist victory in this country in 1918 ushered in a program of socialization which proved a failure. The dominant branch of the socialist movement in Austria is now more aptly called progressive than radical. Hungary experienced a soviet régime for six months in 1919. During the brief period rapid strides were made in nationalizing land and industry, but the whole movement collapsed in the face of the counter revolution in August of that year. The Communist Party has been driven under cover, and the more conservative Social Democrats now hold but a small minority of the seats in the Assembly. In Czechoslovakia a progressive republican government has held power since 1920. The Social Democratic Party holds about a third of the seats in the Assembly, while the Communists constitute an inconsiderable minority. These are indications of the relative strength of socialism throughout Central and Southeastern Europe. The situation may be generalized by saying that the dominant branch of the socialist movement is Marxian, with a cautious, modified program of collectivism, that this socialistic group holds the balance of power in most of these countries, and that there is a strong undercurrent of more extreme radicalism throughout the entire area.

In Western Europe. In France the socialist movement is weakened by its division into a large number of factions, ranging in program from temperate constitutional collectivism to extreme

syndicalism. The thoroughgoing syndicalists abstain from all political action, placing their reliance on direct methods. The socialist groups which take part in politics coöperate on the basis of a working agreement amongst themselves. But the coalition is never stable, while the Communists adopt obstructive tactics which place them in opposition to all government programs. Despite these weaknesses however the combined socialist groups were in 1925 in command in the Chamber of Deputies. France has had a number of socialist ministers in recent years, but they have always failed to win the united support of their followers. The movement is especially strong in the municipal governments, many important cities being completely dominated by the socialists.

In Belgium the last election of 1925 returned the socialists with a larger vote and a stronger position in the Legislature than any other party. Though they are still in the minority, no government can now be formed without their aid. The dominant type of socialism in this country is collectivist. In Italy and Spain the socialist movement, strongly syndicalist in its character, has been suppressed at least temporarily by the conservative dictatorships which have held power during the past few years.

In Great Britain. The growth of the British Labor Party was summarized in another place (Chapter LIII). In the election of 1925 this party lost its control of the government. It polled however five and a half million votes and thereafter held the position of official opposition to the existing government. The program of this party is largely opportunistic, though its basic policies — nationalization of land, mines, and railroads, forcible redistribution of incomes through progressive taxation and public provisions for the poor, stringent government control of enterprise in the interests of labor — are collectivist in principle. More radical socialist programs have never won much acceptance in Great Britain.

In the United States. The professed socialists here are a relatively insignificant fraction of the population. The predominance of agriculture with its system of small independent proprietorship, the relative comfort of all social classes, above all the comparative

freedom of economic opportunity, which fosters an individualistic character in the people, have been obstacles to the growth of socialism here. In national politics the socialists do not number more than one million voters. Their strength is concentrated in certain urban centres, where they have had some success in winning municipal elections. The dominant branch of the American socialist movement accepts the doctrines of Marx and a program of gradually increasing collectivism. There is a noisy communist faction, which preaches violent revolution but which comprises all told a mere handful. The most striking manifestation of social reform in our country in recent years has taken the form of an agrarian movement among the farmers of the Middle West. But this has been due to temporary economic difficulties resulting from the World War. The movement will probably die out, as these difficulties are removed by natural economic forces, without contributing permanent strength to the socialist program.

Superficial criticism of socialism. In seeking to pass judgment upon socialism, we must discount at the outset many of the conventional attacks made upon it. Every reform movement will arouse the antagonism of the conservative forces which always gather about established institutions. All such movements will appear to conflict with prevailing moral standards and will be condemned as heretical, unethical, morally wrong. The present capitalist system was thus condemned in the early stages of its development; so were the modern democratic movement in government, the reform movements in the churches, and other tendencies which are now accepted as parts of the established order of things. Socialism has been passing through this stage of its development. It has been condemned as destructive of moral standards, as atheistic and hostile to religion, as subversive of family relations and the purity of home life, as legalizing theft and violence against persons, as unpatriotic. All of these charges can be substantiated in the case of individual leaders or particular factions of socialism. But they are also true of individuals and groups within the ranks of the conservatives. No strong case against socialism can be established on such grounds. The movement is one of economic

reform and must be judged by its probable effect upon humanity in terms of economic welfare.

Judgment of socialism. The case in favor of socialism may be inferred from its indictment of the capitalist order, which was summarized earlier in this chapter. Every thinking man admits that each of the charges in this indictment contains elements of truth. The present order is no more perfect than is any other work of man. On the contrary it contains many and serious faults, faults of inefficiency, of injustice, above all of retarded and imperfect adjustment to the changing conditions of human life. The socialist movement is to be credited with the great service of having focused attention on these faults and accelerated their correction. Its own program however is not convincing. We may summarize the case against it under four principal heads.

(1) Socialism exaggerates the inefficiency of the present order. The indictment of capitalism is overdrawn. After all has been said that can be said against it, our system has succeeded in maintaining the largest population the world has ever known upon the highest level of average comfort. It has crowded into the short space of a century and a half greater advance in all the arts of production than was achieved during the whole preceding space of society's history. It has equipped mankind with an unprecedented supply of capital instruments and is increasing the quantity and improving the quality of this equipment without pause. The uncontrolled individualistic system of production and marketing, despite occasional lapses, works smoothly and effectively to supply the individual consumer with want-satisfying goods drawn from the resources of the entire world. It has accomplished this by providing incentives which call into play two powerful human impulses: acquisitiveness and the desire to excel in competition. The institution of private property supplies the former of these incentives, individual enterprise, the latter. Socialism aims to suppress the motive of acquisition entirely and to deprive success in competition of all material rewards. In their place it would rely upon the motives of brotherhood, social service, and altruism. It is not necessary to deny that human nature possesses these latter charac-

teristics nor that they are of higher order than those nourished by a capitalist environment. But there is grave reason to doubt that they will prove as effective in stimulating human energy along lines of economic activity as the motives which have been responsible for the unparalleled material success of capitalism.

(2) Socialism underestimates the vitality of capitalism. The survival of social institutions depends upon their power of adjustment to new conditions as they arise and of providing within themselves mechanisms for solving the problems which are created by their own growth. Capitalism has this power. It is a flexible system. The institutions of property and competition can be qualified when need arises without destroying their fundamental utility. Collective bargaining can be substituted for the individual contract when the latter works injury to the weaker members of society. Social insurance can be developed to safeguard the workman against many of the hazards of the wage-earner's life. Labor legislation can assure him minimum standards of safety, sanitation, even of income. Collective provision may be made for the services which private initiative will not supply, such as education, agencies of public health, recreation facilities. Producers can be restrained from undermining and destroying each other by unfair competitive practices. The consumer can be protected against adulterated goods and scrimped quantities by laws enforcing standards of purity and honesty of weight and measure. The overweening power of great business can be restrained by public control. All these things can be done without discarding the basic institutions of society; they are now being done by continuous process. The true conservatives of our day are those who lend their support to these reforms; those who oppose timely readjustment of traditional ways and methods must be ranked among the enemies of the present order.

(3) The socialist program creates insuperable practical difficulties. In testing the practicability of socialism we confront the fact that the socialist programs are diverse and that each school has its own practical problems. Since we cannot examine each of these programs in detail, we shall merely indicate difficulties which are com-

mon to them all. All forms of socialism impose on some public body the task of determining the prices of different goods and services. Is not this a task too great for human intelligence in so complex a society as ours? Price fixing involves a determination of quantities to be produced and a rationing of individual consumption. After wages have been fixed for all classes of labor, the state or some other public body must then see to it that society is supplied with the right amounts of each type of worker. When a price has been set upon wheat or coal or shoes, some authority must then supervise production and consumption. These difficulties multiply as one ponders the problem of authoritative control.

(2) Again some method must be devised for detecting the natural leaders of men, advancing them to the positions for which they are best qualified, and giving them the requisite authority to make their powers effective. Each of these steps creates serious difficulties. For the selection of leaders socialism has to propose no better method than that of public competitive examination. This always works against the innovator in favor of the man who conforms to the traditionalism so dear to the hearts of public officials. It places a handicap in favor of the politician, the time-server, the sycophant. The adjustment of capacity to task is no less difficult. Today it is worked out by process of trial and error. In the society of the socialists it is apt to be decided by rule of thumb. Finally democracy has always been suspicious of concentrated authority. Would discipline be improved if each manager owed his tenure of office to the goodwill of his subordinates? It is not enough for socialists to argue that these problems are not now perfectly solved. They must propose a superior method.

(3) Thirdly socialism must provide for the maintenance and improvement of the capital equipment. No possibility would remain for the saving of capital by individual initiative, since interest would be destroyed as a private income. Now it is not impossible for society to provide a different method of saving. The state, as recipient of the social income, could withhold a portion to be reinvested, thus forcing the people to save against their will or with-

out their knowledge. But, in view of the impatience of men for immediate income and of the subservience of office holders to popular clamor, is it likely that any group of public officials could be found with sufficient courage and foresight to administer such a system wisely? As regards the improvement of technical processes, it is argued that these come in final analysis from scientists who labor without thought of material reward for the sake of their love of truth. This may be granted without removing the difficulty. Today the discoveries of science are adapted to the economic service of mankind, not by the discoverers, but by other men who are certainly intent on personal reward. Pure science might flourish under socialism, but invention would lag behind because of the apathy, the obtuseness, or the abhorrence of change of public officials. → lack of personal reward.

(4)

Finally there is the menace of over-population. Today restraint of the birth rate is based on individual motives, the desire for social advancement, the wish to escape the inconvenience of children, ambition for the welfare of the next generation. In all forms of socialism the children would cease to be a direct burden on their parents. The future of all families would be equally secure regardless of their relative size. Is it not probable that such provision would release the check on population increase and bring into play the first law of Malthus? We do not argue that this danger is inevitably present in socialism, for it cannot be proved that appeals to other motives might not be effective in restraining births. But at any rate it is a menace which may be reasonably expected.

(4) Socialism would destroy human liberty. As was pointed out above, the socialists attack the present order on the ground that it insures no real freedom for the average man. But is it not probable that their system would make this matter worse? It is difficult for us to imagine the extent to which the actions of the individual would be controlled by public authority under a system of socialism, but a faint image of this condition can be had by recalling the situation in all belligerent countries during the World War. A system of public control, when instituted over the major economic

activities of the citizens, ramifies until it embraces practically all the interests of a normal human life. This menace is greatest in the program of state socialism, least so perhaps in that of the guild socialists. But all socialist programs rely, in the final instance, for their power to control the conduct of individuals upon the *command.* In our order compulsion is mixed with *reward* as a director of human energy. To most men a society where the type of work in which a man may engage, the income which he receives, and the kinds of wants he is permitted to satisfy are all determined by authority bears little resemblance to a condition of freedom. In a more fundamental sense it may be urged that such an ordered state of existence, if it succeeded in maintaining itself, would spell stagnation. Progress in society, as in organic evolution, proceeds by differentiation, not by conformity to rule.

Suggestions for Further Reading on Part VIII. Reform of the Economic System

I. Profit Sharing, Co-Partnership, and Coöperation

Burritt, A. W., *et al.* *Profit Sharing.* New York, 1918
Bowie, J. A. *Sharing Profits with Employees.* London, 1922
Emmett, B. *Profit Sharing in the United States.* Bulletin 203, United States Bureau of Labor Statistics
Williams, A. *Co-Partnership and Profit Sharing.* New York, 1913
Heilman, R. E. (Editor) *Profit Sharing.* New York, 1918
Fay, C. R. *Coöperation at Home and Abroad.* New York, 1920
Report on Profit Sharing and Labour Co-Partnership in Great Britain, Cd. 544, 1920
Wolf, H. W. *People's Banks.* London, 1910
Gilman, N. P. *Methods of Industrial Peace: a Dividend to Labor.* New York, 1879
Carver, T. N. *Present Economic Revolution in the United States.* Boston, 1925

II. Socialism

Hillquit, M. *Socialism in Theory and Practice.* New York, 1909
Russell, B. *Proposed Roads to Freedom.* New York, 1919
Sellars, R. W. *The Next Step in Democracy.* New York, 1916
Skelton, O. D. *Socialism: a Critical Analysis.* Boston, 1911
Withers, H. *The Case for Capitalism.* New York, 1920
Tawney, R. H. *The Acquisitive Society.* New York, 1920

WEBB, B. and S. *Decay of Capitalist Civilization.* New York, 1923
WEBB, S. *Socialism and Individualism.* London, 1910
COLE, G. D. H. *Guild Socialism.* London, 1920
FARBMAN, M. S. *Bolshevism in Retreat.* London, 1923
PASVOLSKY, L. *The Economics of Communism.* New York, 1921
HILLQUIT, M. *History of Socialism in the United States.* New York, 1910
WELLS, H. G. *New Worlds for Old.* New York, 1908
BELLAMY, E. *Equality.* New York, 1899

INDEX

Acceptance, i, 417
Advertising, in marketing, i, 151, 156, 157; competitive, ii, 625
Agriculture, unprofitable field for large-scale methods, i, 128, 326, ii, 51; subject to law of increasing profits, i, 332; economic rent in, ii, 97–133; diminishing returns in, ii, 101–109
Amana Society, i, 75
American Federation of Labor, structure of, ii, 544, 545
Anti-trust laws, ii, 57–66
Apportionment, of general property tax, ii, 385
Arbitration, voluntary, ii, 578–81; compulsory, ii, 584–90
Aristotle, on interest, ii, 198
Assembling, a function of marketing, i, 149
Assessment, of general property tax, ii, 383–85, 389–95; results of, ii, 395; breakdown of, ii, 397–99
Assets, defined, i, 104; corporation, i, 105
Australasia, compulsory arbitration in, ii, 585, 587

Babson's Statistical Organization, i, 222
Balance of payments, international, i, 551–65
Balance sheet, a property account, i, 104; corporation, i, 105; bank, i, 418–20, 423, 431–33; international, of U. S., i, 561
Bank deposit, defined, i, 420
Bank notes, i, 405–406, 415, 419, 421, 429; issue of, i, 433; defined, i, 434; as money, i, 435; restrictions on issue of, i, 435; elasticity supplied to currency by, i, 438; of France, i, 447; of England, i, 449–52; of Canada, i, 453; national, i, 457; state, i, 457; local, i, 456, 457; federal reserve, i, 468; national and federal reserve, criticized, i, 474
Bank of England, i, 448–53
Bank of France, i, 447
Bank of the United States, First, i, 455; Second, i, 456
Banking, principles of, i, 412–45; normal business of, i, 428; national systems, i, 446–74; history of American, i, 455–73; wild-cat, i, 456, 457
Banks, investment, i, 113–16, 118; kinds of, i, 414; collection between, i, 424; clearing house an association of, i, 425; private, i, 442, 458; savings, i, 442, 458; central and branch, i, 446; state, i, 458; and the foreign exchange market, i, 538; Morris plan, ii, 180
Bargain sale, i, 226–27, 310
Bargaining, collective, ii, 556–60, 562, 563
Barter, defined, i, 23; money enables escape from, i, 347
Bill of exchange, defined, i, 535; use of, in foreign trade, i, 536, 537; specimens, i, 535, 537; time, specimen, i, 547
Bimetallism, i, 385–97
Birth control, suggested by Malthus, ii, 300; popularized by proscription, ii, 311, 313
Birth rate, ii, 311; causes of declining, ii, 312; differential, ii, 314
Bonds, defined as a liability, i, 105; government, artificial value of, i, 460; government, high standing of, ii, 358
Borrowing, for production, ii, 177–79, 184; for consumption, ii, 177, 180, 182, 185; effect of time preference

on, ii, 180–82; of governments, ii, 182; *vs.* lending, ii, 185; for investment, ii, 195; government, ii, 354–61
Bounties, protection by, ii, 452
Bowley, A. L., on distribution of British income, ii, 291
Brokers, investment, i, 113–19; note, i, 431, 439–41
Brook Farm, i, 75
Brookmire Economic Service, i, 222
Budget, government, ii, 363
Bureau of Labor Statistics, index numbers of, i, 488–91; on profit sharing in U. S., ii, 599
Burritt, A. W., on pseudo profit sharing, ii, 596
Business cycle, i, 502–519; motivating forces of, i, 515; fluctuations of, injurious to all, i, 515; remedies for fluctuations of, i, 518–19; unemployment resulting from, ii, 486, 493; and union membership, ii, 550; socialists on cause of, ii, 626
Buying, rational, i, 309–313
By-products, made useful by large-scale production, i, 124, 132

Canada, banking system of, i, 453–55; compulsory investigation of labor disputes in, ii, 582–84
Capital, a factor of production, i, 29, 32–41, 56; accumulation of, contributed to by present system, i, 207; the means to future income, i, 355; value of, i, 356, 363; property and, i, 357; productivity of, ii, 178–80
Capital stock, explained, i, 101; a liability, i, 105; value of, i, 107; dividends on, i, 109, 110; watered, i, 111; dealings in, i, 114, 116, 291
Capitalism, socialist criticism of, ii, 625–29; socialist underestimate of, ii, 645
Capitalist, no clean-cut distinction from laborer, i, 38; essential service of, i, 39; rise of, i, 55, 68
Carnegie, Andrew, first to use a chemist, i, 126; his use of scrap heap, i, 137
Cattle as money, i, 371, 379
Check, specimens of, i, 422; defined, i, 423; payment by, i, 423–28; limitations of, i, 434; a bill of exchange, i, 535
Chicago Board of Trade, i, 160, 260
Classes, marked off by wages, ii, 232–36; cause of discrepancies in their wages, ii, 237; permanence of cleavage between, ii, 238
Clayton Act, ii, 60; provisions, meant to stifle trusts, conspiracies, and monopolies, ii, 63–65
Clearing house, i, 425–27
Closed shop, ii, 560, 562, 563
Coal, i, 57, 68
Coinage, i, 379; free, defined, i, 379; free, under bimetallism, i, 386, 387, 393, 395
Collectivism, ii, 634
Combination (industrial), i, 128–37
Commercial bank, operations of, i, 414–29
Commission merchant, i, 158
Commodities as money, i, 371, 374
Communism, ii, 638–40
Compensation laws, ii, 532
Competition, desirability of unrestricted, i, 71; in modern industry, i, 90; defined, i, 248; price under, i, 317, 336; cutthroat, i, 336–38, 340, ii, 14; effect on price and industry, i, 343–44; potential, i, 351; railroad, an evil, ii, 5, 14; railroad, maintenance of, ii, 34; where cutthroat, develops, ii, 52; socialism on waste of, ii, 625–26
Competitive price, i, 248–61, 317–19, 336; causes of change in, i, 263–82; laws of, i, 284–94
Conciliation, ii, 571–75
Consumption, growth during Industrial Revolution, i, 69; regulated by price, i, 160; borrowing for, ii, 178, 180, 182; taxes on, ii, 372
Coöperation, ii, 608–22; division of labor a form of, i, 75, 81; unconscious, i, 92; producers', profit sharing a road to, ii, 597, 600; producers', 609–611, 637; credit, ii, 612–14; consumers', 614–22
Co-partnership, ii, 603–08
Copyrights, production aided by, i,

202; system of, a form of monopoly, ii, 49
Corporation, contrasted with partnership, i, 95, 96, 98; described, i, 98–114; utility of, in modern life, i, 111; marketing of its securities, i, 113–19; taxation of, ii, 458–62; and Marxian theory, ii, 633
Cost of production, i, 314–43, 364; analysis of, i, 321; unit, i, 321, 322, 326; marginal and average, i, 328–31; price changes and, i, 341; discounted income and, i, 366; discounted value and marginal, i, 366–67; marginal, in agriculture, ii, 106–11
Costs, increasing, i, 326, 331, 347; constant, i, 326, 327, 328, 333, 349; decreasing, i, 326, 328, 334, 348; decreasing, an unstable condition, i, 336; fixed or "sunk," i, 337; decreasing, may lead to monopoly, i, 340; law of decreasing, railroads an extreme example of, ii, 6, 14; joint, defined, ii, 11
Credit, in marketing, i, 152, 154, 156; commercial, i, 412–13; bank, i, 414, 421–22; bank, as a medium of exchange, i, 436; bank, elasticity of, i, 438–39; function of, i, 443; a solvent factor, i, 443; importance of, i, 444; bank, in period of prosperity, i, 505; control of, as remedy for business fluctuations, i, 518
Credit money, i, 382, 383, 401, 405–407, 408, 409; today, i, 411
Credit system, the essence of, i, 412–13; extent of, determined by savings, i, 413; gain from, i, 443
Crisis, period in business cycle, i, 504, 508–513, 516; panic frequently consequent to, i, 512
Cultivation, margin of: illustration, ii, 114–17; defined, ii, 117; intensive and extensive, ii, 117–21
Currency, stable, importance of, i, 376; classification of, i, 437; elastic, need of, i, 437, 462; quantity of, its effect on value of money, i, 479, 481

Davie, M. R., on differential birth rate, ii, 316
Death rate, ii, 308
Debts, government, ii, 356–61
Demand, i, 227–36; defined, i, 234, 255–57, 284; law of, i, 235, 284, 312; distinguished from desire, i, 260; effect of increase in, i, 263, 275, 277; effect of decrease in, i, 266, 274, 276; what is back of, i, 296, 313; total and individual, i, 297; individual, i, 305–07; influences on, i, 313; unemployment from seasonal fluctuation of, ii, 490–92
Deposit, loan for, i, 419; discount and, i, 421; similarities and differences between note and, i, 434
Depression, period of business cycle, i, 504, 513
Diminishing returns, principle of, in agriculture, ii, 101–12; principle of, another expression of law of increasing marginal cost, ii, 107; social consequences of, ii, 126; from urban lands, ii, 136–39; from capital, ii, 178–80
Discount, process, i, 362; and laws of price, i, 363; bank, defined, i, 414; and deposit, i, 421; changes in bank, i, 432
Discrimination (railroad), between goods, ii, 13, 16; local, ii, 16; local, defense of, ii, 18; local, examples of, ii, 18, 19; local, evils of, ii, 19, 21; personal, ii, 19, 21–23; a cause of government regulation, ii, 27
Distribution, merit in present system of, i, 207; problem of, ii, 97; technical terms of, ii, 98; personal, ii, 273–98; socialist view of, ii, 627
Division of labor, i, 75–83; investment banks an application of, i, 119; waterway development and, i, 170; gain from regional, i, 521–26
Dollar, "stabilized," i, 500; as remedy for business fluctuations, i, 519
Downey, E. H., on work accidents, ii, 528–30
Draft, i, 417; demand or sight, i, 537
Duties, export, ii, 433; import, *see* Tariff; excise, ii, 439, 441; death, ii, 463

Economic rent, in agriculture, ii, 97–133; defined, ii, 100; measure of, ii, 122; on urban lands, ii, 134–45, 153–56; comparative, ii, 139–41; of mines, ii, 145–47; problems of, ii, 147–62; ethical soundness of, ii, 153–62; compared with interest, ii, 165; a cost to the individual, ii, 257; shifting and incidence of tax on, ii, 426
Economics, field of, i, 3–8; defined, i, 8, 12; as a science, i, 9
Education, the source of economic superiority, i, 201
Elasticity, of purchase, i, 236–39, 348, 349, 350; of sale, i, 245; of purchase and excise taxes, ii, 455
Employee representation, ii, 574–76
Entrepreneur, described, i, 40; production controlled by, i, 89; price guides, i, 90, 341; necessary for production, i, 316; limited by competition, i, 318; effects of business cycle on, i, 516; as risk bearers, ii, 76
Environmentalists, on social stratification, ii, 239
Eugenics, ii, 317; negative, ii, 318
Exchange, defined, i, 23; development of, i, 49, 80; function and evolution of, i, 141–47; foreign, i, 373; money a medium of, i, 374; bill of, defined, i, 535; mint par of, i, 540; rate of, i, 541, 544, 549; indirect, i, 544
Exports, financing, i, 535; fluctuating, their effect on foreign exchange, i, 539–40; reaction of exchange rates on, i, 558

Fabian Society, ii, 635
Factory system, rise of organized labor in, ii, 536
Fair, medieval, i, 145
Fairchild, F. R., on Civil War "gold exchange," ii, 90–91
Federal reserve banks, combined statement, i, 473
Federal reserve system, i, 462; organization, i, 463, 465; functions, i, 464–68; reserves of, i, 469; elastic currency furnished by, i, 467, 470, 471
Federal Trade Commission, i, 193, ii, 60; foundation and work of, ii, 61–63
Fees, ii, 351; defined, ii, 352; compared with taxes and government industrial earnings, ii, 352–54
Fiat money, i, 382, 383, 407–408; in history, i, 408, 409; greenbacks as, i, 410; and law of value of money, i, 483, 492–93; collapse of, i, 499; in foreign exchange, i, 549
Fiduciary money, i, 382, 383, 401–407
Fisher, Irving, index numbers of, i, 490; on our fluctuating exports, i, 532
Ford Motor Company, minimum wage policy of, ii, 252; profits of, ii, 262
Foreign exchange, i, 373, 534–50; market, i, 548
Foreign trade, compared with domestic, i, 530; statistics of, i, 531, 532
Franchises, exclusive, a form of monopoly, i, 346, ii, 49
Free trade, in England, i, 68, 86, ii, 439; supported by *laissez faire*, i, 87; *vs.* protection, ii, 441–51

Gambling, the creation of unnecessary risk, ii, 72; compared with investment and speculation, ii, 88
George, Henry, and Single Tax, ii, 158–59
Germany, conciliation of labor disputes in, ii, 573; credit coöperation in, ii, 612–14; socialism in, ii, 640
Ghent system, ii, 504
Gilds, i, 46, 56
Godin Familistère, a pioneer in profit sharing, ii, 595, 599
Gold, why generally employed as money, i, 374–77; as standard money, i, 383; certificates, i, 402; world's production of, i, 494–95; inflation through, i, 494–96; limits to flow of, i, 555–58; effect of its production on trade, i, 556; and national wealth, ii, 443–45
Gold points, i, 542
Gold standard, i, 385; England traditional country of, i, 394; adopted by Germany and France, i, 396; imperfections of, i, 500; possibility of a more stable standard than, i, 500

Government, economic functions of, i, 185–209; control of money, i, 377; war the principal item of expenditure of, ii, 342; state and local, expenditures by, ii, 343; growth of expenditures in, ii, 344; total cost of, in U. S., ii, 346–47; revenues of, ii, 347–48, 350–54; financing primary functions of, ii, 349; borrowing by, 354–61
Government industry, ii, 469–80; finances of, ii, 348
Government ownership, ii, 634; of railroads, ii, 39; in U. S., ii, 470–75; pro and con, ii, 475–80
Government regulation, in English history, i, 83–87; of marketing, i, 161; of industry, i, 192, 208; of prices, i, 223, 225; of monopolies, i, 352, 354; of railroads, ii, 26–45; of trusts, ii, 57–67; of interest rate, ii, 199–201; of business, fees for, ii, 351; by taxation, ii, 378; *vs.* government ownership of public utilities, ii, 479; of child labor, ii, 519–21; of woman labor, ii, 526
Granger laws, ii, 27
Great Britain, distribution of income in, ii, 291; unemployment insurance in, ii, 505–508; legal regulation of child labor in, ii, 519; labor movement in, ii, 537–41; strikes in, ii, 569; conciliation of labor disputes in, ii, 572; mediation of labor disputes in, ii, 577; consumers' coöperation in, ii, 614–21; socialism in, ii, 642
Greenbacks, i, 406; origin and history of, i, 409–411; verification by, of law of value of money, i, 492–93
Gresham's law, stated, i, 388; action of, in Civil War, i, 393; working of, 407, 483
Guild socialism, ii, 636

Hadley, A. T., on local discrimination by railroads, ii, 19
Handicraft Stage, i, 45–51
Harvard Committee on Economic Research, i, 149, 222
Hedging, ii, 84–87; defined, ii, 85
Immigration, i, 199, ii, 319–36
Imperialism, cause of, ii, 128
Imports, the aim and end of commerce, i, 525; financing, i, 537; fluctuating, their effect on foreign exchange, i, 539–40; effect of exchange rates on, i, 558
Incidence, defined, ii, 377; of taxes on property and income, ii, 426–32; of customs tariff, ii, 434, 437, 441; of excise taxes, ii, 454
Income, defined, i, 17; net, i, 17; distinguished from wealth, i, 22; present value of future, i, 355–67; relation to capital, i, 362; discounted, and cost of production, i, 366; greater utility of an even, i, 516; interest rate a factor in valuations of, ii, 188; personal distribution of, ii, 273–98; national and per capita, ii, 273–76; national, proportion of government expenditures to, ii, 341; earned, tax on, ii, 412; corporation, tax on, ii, 413; exempt from taxation, ii, 415, 417, 421; shifting and incidence of taxes on, ii, 426–32
Income tax, ii, 408–22; proportional, progressive, and regressive, ii, 370–72; exemption from, ii, 371; and general property tax, ii, 405; rates, ii, 411, 422–23; administration of, ii, 424–25; state, ii, 425; shifting and incidence of, ii, 431; corporations and, ii, 458
Index numbers, i, 483–90, 509, 511; defined, i, 485
Individual initiative, principle of, i, 88
Individual proprietorship, i, 94
Industrial accident, ii, 526–34
Industrial Revolution, i, 55–74; large-scale production a result of, i, 120
Industry, household system of, i, 43–45; Handicraft Stage of, i, 45–51; domestic system of, i, 51–53, 55, 71; factory system of, i, 53, 55, 60–64, 73; national facilities for, i, 202; diversity of, promoted by protective tariff, ii, 442, 445; infant, protection of, ii, 449
Inequality, universal, ii, 279–80;

problems raised by, ii, 289–98; in taxation, ii, 398
Inflation, through silver, i, 398–400; through gold, i, 494; in World War, i, 496–97; in German currency, i, 499; rise in price level caused by, i, 502
Injunction, failure of, against strikes, ii, 571
Insolvency, i, 107
Insurance, a means of transferring risk, ii, 78–80; industrial accident, ii, 533
Interest, defined, i, 360; nature of, ii, 163–69; social problem of, ii, 196–201; a cost in any business, ii, 257
Interest rate, ii, 183–96; government regulation of, ii, 199–201
"Interests," the, i, 352
Intermediate goods, i, 7, 33, 297
Interstate Commerce Act (1887), ii, 29, 30–32
Interstate Commerce Commission, i, 193, ii, 29, 30, 31–36; given almost unlimited powers of regulation, ii, 40–45
Invention, i, 38, 55, 57, 59, 60, 61, 167, 171; unemployment from progress of, ii, 487; effect of, on hours of labor, ii, 515
Investigation, compulsory, of labor disputes, ii, 582–84
Investment, made easy by corporations, i, 112, 113
Iron, i, 57, 64, 68

Jones, Eliot, on U. S. Steel Corporation, i, 135; on railroad costs, ii, 8, 9–10

Kansas, compulsory arbitration in, ii, 586
Knights of Labor, ii, 543–44

Labor, a factor of production, i, 29, 30–32; immobility of, an obstacle to trade and price equilibrium, i, 529; provisions concerning, in railroad legislation, ii, 41–43; and the interest rate, ii, 194; demand and supply of, ii, 203–211; diminishing productivity as applied to, ii, 205; law of demand for, ii, 207; classifications of, ii, 231; "foreign pauper," ii, 447; defensive policies of, against unemployment, ii, 494–97; adjusting supply of, ii, 497–501; stabilizing the demand for, ii, 501–503; hours of, ii, 511–16; modern child and woman, ii, 516–26; monopoly of, sought by trade unions, ii, 563
Labor exchanges, ii, 499–501
Labor Party, British, growth in power of, ii, 540; leadership of, ii, 636; program of, ii, 642
Labor troubles, i, 63, 72, 131; in period of prosperity, i, 506; of railroads, vitally important, ii, 5
Laborer, no clean-cut distinction from capitalist, i, 38; effect on, of Industrial Revolution, i, 70; chief victim of fluctuations in business activity, i, 516, 517
Laissez faire, explained, i, 71, 86; present system a material departure from, i, 208; government policy of, ii, 548
Land, a factor of production, i, 29; and capital, ii, 163; significance of interest rate in valuation of, ii, 189; shifting and incidence of tax on value of, ii, 428–29
Large-scale production, i, 112, 120–28, 346
Legal tender, defined, i, 381; representative money normally not, i, 401
Liabilities, defined, i, 104; corporation, i, 105
Limited liability, i, 96, 98, 100, 111
Limited partnership, i, 97
"Limping standard," i, 397, 400
Liquidation, period in business cycle, i, 504, 508–513
Loan sharks, ii, 200
Loans, i, 418–19, 421; paying, i, 428; international, effect on balance of payments, i, 559; supply of, ii, 169–76; demand for, ii, 177–83; government, ii, 354–61
Lockout, ii, 564
Losses, business, ii, 262–64

Maison Léclaire, a pioneer in profit sharing, ii, 595, 599, 601

Maisons du Peuple, ii, 621
Malthus on population, ii, 299; criticism of, ii, 300; modified statement of theory, ii, 301
Market, defined, i, 228; perfect, i, 285, 291
Market ratio, *vs.* mint ratio, i, 386–90, 391–97
Marketing, organization of, i, 139–63
Marx, Karl, doctrines of, ii, 629–34
Mediation, ii, 576–78
Mercantilism, i, 86; fallacies of, i, 551; in arguments for protection, ii, 443, 444
Merchant marine, i, 181–84
Middlemen, services of, in marketing, i, 160
Mill, J. S., on increase of population, ii, 301
Mines, economic rent of, ii, 145–47
Minimum wage laws, effect of their adoption, ii, 217, 253
Mint ratio, i, 386–90, 391–97
Monetary system, essentials of modern, i, 378–411; elasticity of British, i, 451; United States, i, 474
Money, i, 371–411; definition, functions, qualities of, i, 373, 374; "lawful," i, 406; history of American, i, 408–11, 492–97; U. S., circulation statement of, i, 475; value of, i, 476–501; value of, defined, i, 478; law of value of, i, 482, 490–98; far-reaching effects of fluctuations in, i, 498
Monopolies, dependent on corporate form, i, 113; growth of, i, 129; regulated by government, i, 193, 352; through cutthroat competition, i, 340; generally not in public interest, i, 353; classification and description of, ii, 46–57
Monopoly, prices, i, 345–54; railroads, ii, 24–25; industrial, ii, 46–57; future of, ii, 67–70; profits and, ii, 269; private, *vs.* government ownership, ii, 478–80; Karl Marx on, ii, 632
Mores, of high economic significance, i, 198–99
National banks, inaugurated, i, 457; notes of, i, 457; notes, security for, destroys elasticity, i, 458–59
National (U. S.) debt, history of, ii, 359
National Industrial Conference Board, figures of government expenditures, ii, 341, 346; figures of total and per capita taxes, ii, 365
Nature and nurture theories, ii, 239
Negligence, doctrine of, ii, 531, 532; contributory, ii, 531; of the fellow servant, ii, 532
Nelson, N. O., Company, profit sharing by, ii, 598
Neo-Malthusianism, ii, 313
New York Stock Exchange, i, 117, 261, 291
Non-Partisan League, ii, 643
Notes, promissory, i, 414–17, 421; unsecured, i, 416; secured, i, 416; endorsed, i, 416; federal reserve, i, 467
Nurture. *See* Nature and nurture theories

Oneida Community, i, 75
Over-population, Malthus' forebodings of, ii, 299; still a menace, ii, 303, 647; *vs.* underpopulation, ii, 319–21, 326
Ownership, absentee, ii, 154

Panic, i, 320, 463; of 1814, i, 456; of 1837, i, 456; old-time, now vanquished, i, 471; prevented in 1920, i, 513; of 1907, i, 512; of 1837, government surplus partly responsible for, ii, 362; of 1837 and local craft unions, ii, 541
Paper, commercial, i, 414–17; "one-name," "two-name," "three-name," i, 416, 447; commercial, houses, i, 440
Paper money, i, 380
Partnership, i, 95–97. *See also* Limited partnership
Patents, production aided by, i, 202; system of, a form of monopoly, ii, 48
Personal liberty, principle of, i, 88
Pigou, A. C., on personal ability and

inequality, ii, 293; on mediation of labor disputes, ii, 576–77
Pool, a monopoly device, ii, 53; the Wire Nail, ii, 54
Population, ii, 299–318; spread of, ii, 319
Poverty, problem of, ii, 286
Power, i, 60–63, 76; machinery and, i, 123
Price, a fundamental concept, i, 22; defined, i, 24; production and consumption regulated by, i, 90–92, 160; significance of, i, 217–39; market, i, 250–54, 285; fixed by cost of production, i, 254, 317; fixed by sole buyer, i, 254; causes of changes in, i, 263–82; supply and, i, 331; monopoly, i, 345–54; higgling, i, 345; discount and laws of, i, 363; money and, i, 476–501; marginal cost equated to market, ii, 109; not governed by economic rent, ii, 141–45; not determined by profits, ii, 270
Price equilibrium, law of, i, 528; obstacles to, i, 529
Price level, general, law of, i, 482, 490–98; and business cycle, i, 502
Private property, principle of, i, 88; saving stimulated in system of, i, 207; under socialism, ii, 624; socialist view of, ii, 626
Produce exchanges, i, 159
Production, i, 26–41, 68; division of labor increases, i, 77–79; control of, i, 83–87, 90–92; motive to, i, 89; corporation an important factor in, i, 112; modern specialized, i, 140; aided by patents and copyrights, i, 202; cost of, i, 314–43; competitive, i, 317; monopoly, i, 345; local advantages determine kind of, i, 521; regional specialization an aid to, i, 522; law of comparative advantages in, i, 525; agricultural, regulated by economic rent, ii, 124; economic rent not a cost of, ii, 129–33; borrowing for, ii, 177–80; labor cost of, ii, 447; effect of excise taxes on, ii, 455–56; under socialism, ii, 624, 627
Profit sharing, ii, 595–603, 608
Profits, ii, 256–72; defined, ii, 260
Progress and Poverty, ii, 158
Prohibition, and the excise system, ii, 453
Property, defined, i, 18; in free persons, i, 20; distinguished from wealth, i, 20; evidence of, i, 21; shifting and incidence of taxes on, ii, 426–32
Property account, a balance sheet, i, 104; of a corporation, explained, i, 105
Property tax, general, ii, 381–404; technical terms relating to, ii, 388; general, and income tax, ii, 405; general, on corporations, ii, 459
Proportionality, law of, stated, ii, 243; applied to labor groups, ii, 246
Prosperity, period of business cycle, i, 504–508
Protection, defined, ii, 437; *vs.* free trade, ii, 441–51; analysis of arguments for, ii, 443–51; urged from military necessity, ii, 451. *See also* Protective tariff
Protective tariff, ii, 436, 439–52; abolished in England, i, 68; opposed by *laissez faire*, i, 87; enactment by U. S. a stimulus to manufacture, i, 171; purpose of, i, 344, ii, 378. *See also* Protection
Public debt, retirement of, i, 459
Public finance, ii, 341–64
Public utilities, ii, 5; generally natural monopolies, ii, 50; government ownership *vs.* government regulation of, ii, 479

Race suicide, from differential birth rate, ii, 316, 326
Raiffeisen system, ii, 612–14
Railroads, first, in England, i, 66–68; history of American, i, 173–78; capitalization of, i, 178; economic services of, i, 178; interdependence of public and, ii, 3–5; in class of public utilities, ii, 5; competition among, an evil, ii, 5, 14; labor problems of, ii, 5; subject to law of decreasing costs, ii, 6, 8, 14;

fixed costs, ii, 6–8; operating costs, ii, 7; joint costs, ii, 10–11; rates of, ii, 11–24; breakdown of competition between, ii, 23; monopoly in, inevitable and economical, ii, 24–25; government regulation of, ii, 25, 26–45; government operation, during World War, ii, 36–38
Recovery, period in business cycle, i, 514
Rediscount, i, 431
Rent, defined, ii, 98. *See also* Economic rent
Representative money, i, 382, 383, 401–403
Reserve, bank, i, 429; methods of controlling, i, 430–33; influence of, on credit, i, 439; to reduce risk, ii, 75
Retail stores, i, 152–55
Returns, diminishing, law of, i, 327
Revenue, history of U. S. system of, ii, 405–11
Ripley, W. Z., on local rate discrimination by railroads, ii, 17
Risk, ii, 71–91; in marketing, i, 151, defined, ii, 71; reduction of, ii, 73–76, 84–87; law of large numbers (stated) and, ii, 74; transferring, ii, 76–84; profits, insurance, and, ii, 264–69; doctrine of assumption of, ii, 532
Rochdale system, ii, 615
Roosevelt, President, his campaign against trusts, ii, 59–60
Ross, E. A., on comparative death rates, ii, 310
Royalty, a kind of economic rent, ii, 146
Russia, communism in, ii, 638–39

Salaries, as prices, i, 219
Saving, i, 35–37; strong motives for, in system of private property, i, 207; volume of, determines extent of credit system, i, 413; individual, ii, 173; involuntary, ii, 174; corporate, ii, 174; by insurance companies, ii, 176
Schulze-Delitsch system, ii, 612–14
Securities, tax-free, ii, 417–18; tax-free, loss and gain from, ii, 418–20; exemption of, ii, 462
Security holding companies, ii, 56, 57
Sherman Anti-Trust Act (1890), ii, 35, 58–60, 62, 63
Silver, why generally adopted as money, i, 374–77; free coinage of, abolished, i, 393, 396; inflation brought by Bland-Allison and Sherman acts for coining, i, 398–400; certificates, i, 402
Silver standard, i, 385
Single Tax, ii, 157–62
Socialism, ii, 623–48; compulsory arbitration the "opening wedge" of, ii, 589; defined, ii, 625
Specialization, i, 46, 49; furthered by combination, i, 131; in marketing, i, 162, 163; regional, an aid to production, i, 522, 524; limited by law of diminishing returns, i, 527. *See* Division of labor
Speculation, defined, ii, 80; analysis of, ii, 80–88; compared with investment and gambling, ii, 88; judgment of, ii, 89
Standard money, i, 382; gold coin the, i, 383; essentials of, i, 384
Standard of living, i, 45, 63, 70, 72, 73; theory of wages, ii, 248; true effect of, ii, 249; minimum, ii, 287
Statement, corporation, i, 105; bank, i, 418–20, 423, 431–33
Stock. *See* Capital stock; Supply
Stock dividends, i, 110
Stock exchanges, i, 116; "booms" and "panics" on, i, 221; in period of prosperity, i, 507; importance of speculator on, ii, 81–82, 84
Stoppage at source, ii, 459
Strikes, i, 63, 131, 150, ii, 564–68; an insignificant cause of unemployment, ii, 487; devices to prevent or settle, ii, 569–91
Subsidies, ship, ii, 452
Substitutes, monopoly price limited by, i, 350
Supply, i, 241–46; defined, i, 243, 258–59, 284; law of, i, 245, 284; distinguished from stock, i, 260; effect of increase in, i, 268, 276, 277; effect of decrease in, i, 271, 274, 275, 277; controlled by cost

of production, i, 321; price and, i, 331
Surplus, producer's, i, 46, 80, 82, 91; buyer's, i, 308; producer's, in agriculture, ii, 110–13; government, an embarrassing problem, ii, 361
Syndicalism, ii, 636

Taft, President, against trusts, ii, 60
Tariff, customs, ii, 433–52; revenue, ii, 439; protective, ii, 436, 439–52. *See also* Protective tariff
Taussig, F. W., on types of workers, ii, 231–32
Taxation, general principles of, ii, 365–80; technical terms in, ii, 379; direct, constitutional restriction of, ii, 406; indirect, in American history, ii, 407–08; excise, ii, 463; double, ii, 466
Taxes, defined, ii, 350; compared with fees and government industrial earnings, ii, 352–54; proportional, ii, 369; progressive, ii, 370; regressive, ii, 371, 373; consumption, ii, 372; business, ii, 373; poll, ii, 373, 468; escape from, ii, 374; evasion of, ii, 376; shifting of, ii, 376; direct and indirect, ii, 377; shifting and incidence of, ii, 426–32; excise, ii, 439, 441, 453–58; internal revenue, ii, 453; consumption, generally regressive, ii, 457; state excise, ii, 457; corporation, ii, 458–62; inheritance, ii, 463–68. *See also* Income tax
Time preference, the reason for interest, ii, 164–65; and marginal utility, ii, 171; defined, ii, 172; individual, ii, 180
Tobacco as money, i, 371, 372, 373, 375, 376
Token money, i, 382, 383, 400, 401, 403–405
Trade, function of, i, 141; evolution of, i, 141–47; increased with transportation improvements, i, 164–77; volume of, its effect on value of money, i, 479, 481; interregional, i, 520–33; controlled by price, i, 526–28; obstacles to, i, 529; U. S. domestic, i, 530; U. S. foreign, i, 531; "favorable" and "unfavorable" balances of, fallacies in, i, 551, 562–65; favorable balance of, as argument for protection, ii, 443–45
Trade Commission Act, ii, 60, 61–63
Trade union movement, ii, 535–68
Trade unions, contrasted with gilds, i, 47; beginnings of, i, 55, 72, 208; their influence toward uniformity, ii, 220; structure of, ii, 545; American, growth of, ii, 549; benefit systems of, ii, 555; opposition by, to compulsory investigation of disputes, ii, 584; opposition by, to profit sharing, ii, 603
Transportation, i, 43, 65–68, 82
Transportation, a function of marketing, i, 151; organization of, i, 164–84; and law of comparative advantage, i, 527
Transportation Act (1920), ii, 39–43
Treasury, Independent, founded, i, 456; Independent, dangerous system of, modified, i, 472; notes of 1890, i, 399, 473, 474
Trust, built up on basis of corporation, i, 113; examples of, i, 129; dissolution of tobacco, i, 130; production by, i, 324; a result of cutthroat competition, i, 340; unpopularity of, i, 352; defined, ii, 52; technically defined, ii, 55
Trust companies, i, 441, 458
Trusts: Standard Oil, ii, 55, 56, 60, 62, 63; Sugar, ii, 56, 57, 58; Tobacco, ii, 57, 60, 62, 63; government regulation of, ii, 57–67

Under-population *vs.* over-population, ii, 319–21, 326
Underwriting, i, 114; defined, i, 115
Unearned increment, ii, 156–62
Unemployment, ii, 485–510; remedies for, ii, 497–503; insurance against, ii, 503–10, 556
Union, industrial, ii, 545–47; industrial, field for, ii, 554; closed, ii, 560. *See also* Trade unions
United States, industrial development in, i, 72–74
United States notes. *See* Greenbacks

United States Post Office, ii, 471–75; discrimination by, ii, 473, 476
United States Steel Corporation, scope of its operations, i, 133; co-partnership scheme of, ii, 604
Usury, ancient and modern, ii, 199
Utility, i, 11; creation of, i, 26; three forms of, i, 27; time, created by marketing, i, 139; place, created by transportation, i, 139; how it varies, i, 297; subjective side of, i. 298; expressed quantitatively, i, 299; marginal, defined, i, 300; marginal, law of, i, 302; marginal and total, i, 303–13; marginal, i, 315, 344; speculator creates time, ii, 83; law of decreasing, ii, 88; marginal, and time preference, ii, 171

Value, a fundamental concept, i, 22; defined, i, 23; distinguished from price, i, 24, 476; par, i, 108; book, i, 108, 109, 111; market, i, 109; marginal utility and, i, 305; discounted, and marginal cost, i, 366–67; money furnishes unit of, i, 374; labor theory of, ii, 630; surplus, ii, 631

Wage problems, ii, 241–55
Wages, as prices, i, 219; general law of, ii, 202–30, stated, 210; defined, ii, 202; sphere of bargaining in, ii, 214–18; forces toward uniformity in, ii, 220–21; fundamental importance of marginal productivity for, ii, 227; differences in, ii, 231–41; real *vs.* money, ii, 241; standard of living theory of, ii, 248–51; under protective tariff, ii, 446–49; relation of hours to, ii, 513; Marxian law of, ii, 630
Wages of management, from investments, ii, 192; a cost in any business, ii, 258–60
"Walking delegates," ii, 559
Wampum, i, 371, 372, 373, 375
Wants, i, 3; divided into necessities and luxuries, i, 4–7; unsatisfied without effort, i, 7
War, principal source of national debts, ii, 355, 359–61; socialist view of origin of, ii, 626
Watered stock, i, 111; an incentive to corporate saving, ii, 175
Wealth, defined, i, 12; yields both income and costs, i, 17; distinguished from property, i, 20; distinguished from income, i, 22; production of, i, 26–27; government consumption, production, and distribution of, i, 185–86; unequal distribution of wealth, i, 205; as capital, i, 355; personal distribution of, ii, 281–84; national, and gold, ii, 443–45
Webb-Pomerene Act, permitting combinations in export trade, ii, 65
Whitley Industrial Council Plan, ii, 572
Wholesale merchant, i, 155
Wilson, President, against private monopoly, ii, 59; Trade Commission and Clayton acts a result of his recommendations, ii, 60
World War, impairment of national credit through, ii, 360

Zollverein, ii, 451